Math 1554

Milan Capoor

Fall 2021

1 Module 1: Linear Equations

1.1 TOPIC 1: Systems of Linear Equations

Linear equation: an equation in the form $a_1x_1 + a_2x_2 + ... + a_nx_n = b$ where a_n and b are real or complex coefficients of the variables x_n

Linear system (system of linear equations): a collection of linear equations involving the same variables

Solution: a list of values $(s_1, s_2, ..., s_n)$ for which substitution into the variables $(x_1, x_2, ..., x_n)$ makes the equations true - Found where the equations' lines intersect

Solution set: the collection of all solutions to a linear system

Equivalent: characteristic of two linear systems if they have the same solution sets

Consistent system: a linear system which has infinitely many solutions or one solution - Happens when the lines intersect at a single point or when the lines coincide

Inconsistent system: system with no solutions - When the lines are parallel

Matrix: the essential information of a linear system can be recorded compactly in a rectangular array

Given the system

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 \quad - 5x_3 &= 10 \end{aligned}$$

The *coefficient matrix*

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}$$

encodes the essential information of the system while the *augmented matrix* is

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

An $m \times n$ matrix is a rectangular array of numbers with m rows and n columns.

1.1.1 Solving a linear system

The basic strategy is to *replace one system with an equivalent system that is easier to solve.* Three basic operations are used to simplify a linear system:

1. (Replacement) Replace one equation by the sum of itself and a multiple of another equation
2. (Interchange) Swap two equations
3. (Scaling) Multiply all the terms in an equation by a nonzero constant

1.1.2 Example 1: Solve the system given above

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 \quad - 5x_3 &= 10 \end{aligned}$$

Row Reduction Procedure:

1. Construct the augmented matrix

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

2. $R_3 - 5R_1$

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -10 & 10 \end{bmatrix}$$

3. $\frac{1}{2}R_2$

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

4. $R_3 - 10R_2$

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 30 & -30 \end{bmatrix}$$

5. $\frac{1}{30}R_3$

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

6. $R_1 - R_3 and R_2 + 4R_3$

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

7. $R_1 + 2R_2$

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

8. Verify the solution $(1, 0, -1)$ by substituting it for (x_1, x_2, x_3) in the original system

$$\begin{aligned} (1) - 2(0) + (-1) &= 0 \\ 2(0) - 8(-1) &= 8 \\ 5(1) - 5(-1) &= 10 \end{aligned}$$

The equations agree so $(1, 0, -1)$ is indeed a solution to the system. ■

These row operations can be applied to any matrix —not just those that arise as the augmented matrices of a linear system.

Row equivalent: quality of two matrices if there exists a sequence of elementary row operations that transforms one matrix into the other

All row operations are reversible. Hence, *If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.*

1.1.3 Existence and Uniqueness

A large focus of the course and of the analysis of linear systems generally depends on asking two questions:

- Is the system consistent?
- If a solution exists, is it unique?

These questions can often be answered from the triangular form of the matrix (in Example 1 this was $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & -1 \end{bmatrix}$) by understanding that $x_3 = -1$ and solving the other equations from there.

If this method creates contradictions, however, such as in the system $\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 1 \end{bmatrix}$, then the originial system is inconsistent.

1.2 TOPIC 2: Row Reduction and Echelon Forms

Nonzero: any row or column which contains at least one nonzero entry

Leading entry: the leftmost nonzero entry of a row

Row echelon form: a matrix is in row echelon form if it has the following three properties:

1. All nonzero rows are above any rows of all zeros
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it
3. All entries in a column below a leading entry are zeroes

Reduced echelon form: a matrix in row echelon form which satisfies the additional criteria:

4. The leading entry in each nonzero row is 1
5. Each leading 1 is the only nonzero entry in its column

These are the "triangular" matrices of section 1.1.

Each matrix is row equivalent to one and only one reduced echelon matrix.

If a matrix A is row equivalent to an echelon matrix U, we call U an *(reduced) echelon form (REF/RREF) of A*

1.2.1 Pivot Positions

Pivot position: a location in a matrix A that corresponds to a leading 1 in the reduced echelon form of A.

Pivot column: a column of A that contains a pivot position

1.2.2 Example 2: Row reduce matrix A below to echelon form and locate its pivot columns.

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

Solution:

1. Interchange R_1 and R_4

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

2. $R_2 + R_1$ and $R_3 + 2R_1$

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

3. $R_3 - \frac{5}{2}R_2$ and $R_4 + \frac{3}{2}R_2$

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix}$$

4. Interchange R_3 and R_4

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Matrix A is thus in echelon form and so columns 1, 2, and 4 are pivot columns.

Pivot: a nonzero number in a pivot position which is used during row reduction to create zeros

1.2.3 The Row Reduction Algorithm

1. Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

2. Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

3. Use row replacement operations to create zeros in all positions below the pivot.
4. Cover the row containing the pivot position and all rows above it. Apply steps 1-3 to the remaining submatrix, iterating until there are no more nonzero rows to modify. At this time we have the row echelon form.
5. This step is used to find the reduced row echelon form. Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 via scaling operation.

1.2.4 Solutions of Linear Systems

The row reduction algorithm leads directly to the solution set of a system when the algorithm is applied to the augmented matrix of the system.

Basic variables: the variables of a linear system corresponding to pivot columns in the reduced echelon matrix

Free variables: the other variables in the system

Whenever a system is consistent, the solution set can be described explicitly by solving the reduced system of equations for the basic variables in terms of the free variables.
Each different choice of a free variable determines a (different) solution of the system and every solution of the system is determined by a choice of the free variables.

1.2.5 Example 4: Find the general solution of the linear system whose augmented matrix has been reduced to

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

Moving from echelon form to reduced echelon form we have:

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 2 & -8 & 0 & 10 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

Which is equivalent to

$$\begin{aligned} x_1 + 6x_2 + 3x_4 &= 0 \\ x_3 - 4x_4 &= 5 \\ x_5 &= 7 \end{aligned}$$

Because the pivot columns are 1, 3, and 5, the basic variables are $x_1, x_3, and x_5$, leaving x_2 and x_4 as free variables. Thus

$$\begin{cases} x_1 = -6x_2 - 3x_4 \\ x_2 \text{ is free} \\ x_3 = 5 + 4x_4 \\ x_4 \text{ is free} \\ x_5 = 7 \end{cases}$$

1.2.6 Existence and Uniqueness Questions (Redux)

When a system is in echelon form and contains no equation of the form $0 = b (b \neq 0)$, *every* nonzero equation contains a basic variable with a nonzero coefficient. Either the basic variables are completely determined (no free variables), signifying a unique solution, or the basic variables can be expressed in terms of one or more free variables, signifying infinitely many solutions. Thus: **A linear system is consistent iff the rightmost column of the augmented matrix is *not* a pivot column (i.e. there are no rows of the form $\begin{bmatrix} 0 & ... & 0 & b \end{bmatrix}$).**

1.3 TOPIC 3: Vector Equations

Column vector: a matrix with only one column

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

The set of all vectors with two entries is denoted $\mathbb{R}^2$ ("r-two"), where $\mathbb{R}$ refers to the real number entries of the vector and the exponent 2 indicates that each vector contains two entries.
Vectors in $\mathbb{R}$ are ordered pairs which are equal iff their corresponding entries are equal.

Given two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, their sum is

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}$$

while the scalar multiple of a vector is

$$\mathrm{c}\mathbf{u} = \mathrm{c}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} c \cdot u_1 \\ c \cdot u_2 \end{bmatrix}$$

1.3.1 Geometric Descriptions of $\mathbb{R}^2$

Because each point in the cartesian plane is determined by an ordered pair of numbers, we can identify a geometric point with the column vector $\begin{bmatrix} a \\ b \end{bmatrix}$. So we may regard $\mathbb{R}^2$ as the set of all points in the plane.

Parellelogram rule for vector addition: If $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$ are represented as points in the plane, then $\mathbf{u} + \mathbf{v}$ corresponds to the fourth vertex of the parellelogram whose other vertices are $\mathbf{u}$, $\mathbf{0}$, and $\mathbf{v}$.

1.3.2 Algebraic Properties of $\mathbb{R}^n$

For all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{R}^n$ and all scalars c and d:

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

3. $\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$
4. $\mathbf{u}+(-\mathbf{u})=-\mathbf{u}+\mathbf{u}=\mathbf{0}$
5. $c(\mathbf{u}+\mathbf{v}) = \mathrm{c}\mathbf{u} + \mathrm{c}\mathbf{v}$
6. $(c+d)\mathbf{u}=\mathbf{cu}+\mathbf{du}$
7. $c(d\mathbf{u})=(\mathbf{cd})\mathbf{u}$
8. $1\mathbf{u}=\mathbf{u}$

Linear combination: a composite formed of vectors in $\mathbb{R}^n$ and corresponding scalar weights ($\mathbf{y}= c_1\mathbf{v_1} + ...+c_p\mathbf{v_p}$)

1.3.3 Example 5: Solve $x_1\mathbf{a_1}+x_2\mathbf{a_2}=\mathbf{b}$ for $\mathbf{a_1}=\begin{bmatrix}1\\-2\\-5\end{bmatrix}$, $\mathbf{a_2}=\begin{bmatrix}2\\5\\6\end{bmatrix}$, and $\mathbf{b}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$.

We shall begin by constructing an augmented matrix and finding its reduced row echelon form:

$$\begin{bmatrix}1&2&7\\-2&5&4\\-5&6&-3\end{bmatrix} \sim \begin{bmatrix}1&2&7\\0&9&18\\0&16&32\end{bmatrix} \sim \begin{bmatrix}1&2&7\\0&1&2\\0&1&2\end{bmatrix} \sim \begin{bmatrix}1&2&7\\0&1&2\\0&0&0\end{bmatrix}$$

So $x_2=2$ and $x_1=7-2x_2=3$.

Checking: $3\begin{bmatrix}1\\-2\\-5\end{bmatrix}+2\begin{bmatrix}2\\5\\6\end{bmatrix}=\begin{bmatrix}3+4\\-6+10\\-15+12\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$ $\blacksquare$.

A vector $\mathbf{b}$ *can be generated by a linear combination of* $\mathbf{a_1},...,\mathbf{a_n}$ *iff there exists a solution to the linear system corresponding to the augmented matrix* $\begin{bmatrix}\mathbf{a_1} & \mathbf{a_2} & ... & \mathbf{a_n} & \mathbf{b}\end{bmatrix}$. One of the key ideas in linear algebra is to study the set of all vectors that can be written as a linear combination of a fixed set $\{\mathbf{v_1},...,\mathbf{v_p}\}$ of vectors.

The subset of $\mathbb{R}^n$ *spanned by* $\mathbf{v_1},...,\mathbf{v_p}$: (written Span$\{\mathbf{v_1},...,\mathbf{v_p}\}$) is the collection of all vectors that can be written in the form $c_1\mathbf{v_1}+c_2\mathbf{v_2}+...+c_p\mathbf{v_p}$ with $c_1,...,c_p$ scalars.

Asking whether a vector is in Span$\{\mathbf{v_1}, ..., \mathbf{v_p}\}$ amounts to asking whether the linear system with associated augmented matric has a solution.

Note: Span$\{\mathbf{v_1}, ..., \mathbf{v_p}\}$ contains every scalar multiple of $\mathbf{v_1}$ including $\mathbf{0}$

1.3.4 Geometric description of Span$\{\mathbf{v}\}$

For $\mathbf{v} \in \mathbb{R}^{\mathbf{3}}$, Span$\{\mathbf{v}\}$ is the set of points on the line in $\mathbb{R}^3$ through $\mathbf{v}$ and $\mathbf{0}$.
For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, Span$\{\mathbf{u}, \mathbf{v}\}$ is the plane in $\mathbb{R}^3$ that contains $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{0}$.

1.3.5 Example 6: Span$\{\mathbf{a_1}, \mathbf{a_2}\}$ is a plane through the origin in $\mathbb{R}^3$. Is $\mathbf{b}$ in that plane?

$$(\mathbf{a_1} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \mathbf{a_2} = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}, \text{and } \mathbf{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix})$$

This question amounts to asking if the equation $x_1\mathbf{a_1} + x_2\mathbf{a_2} = \mathbf{b}$ has a solution. From row reduction:

$$\begin{bmatrix} 1 & 5 & -3 \\ -2 & -13 & 8 \\ 3 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & -18 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{bmatrix}$$

This suggests that $0 = -2$ but as this is not possible, the matrix is inconsistent and so the vector equation has no solution. Thus, $\mathbf{b}$ is not in Span$\{a_1, a_2\}$. ■

1.4 TOPIC 4: The Matrix Equation

- $\in$: belongs to

- $\mathbb{R}^n$: the set of vectors with n real-valued elements

- $\mathbb{R}^{m \times n}$: the set of real-valued matrices with m as rows and n as columns

Example: The notation $\vec{x} \in \mathbb{R}^5 m$ means that $\vec{x}$ is a vector with five real-valued elements.

1.4.1 Matrix-Vector Product as a Linear Combination

If $A \in \mathbb{R}^{m \times n}$ has columns $\vec{a_1}, ..., \vec{a_n}$, then the matrix vector product $A\vec{x}$ is a linear combination of the columns of A.

$$A\vec{x} = \sum_{n=1}^{n} x_n \vec{a_n}$$

Note that $A\vec{x}$ is in the span of the columns of A.

This means that the solution sets for

$$A\vec{x} = \vec{b}$$

is the same as

$$x_1\vec{a_1} + ... + x_n\vec{a_n} = \vec{b}$$

which is again equivalent to the set of linear equations with the augmented matrix

$$\begin{bmatrix} \vec{a_1} & \vec{a_2} & ... & \vec{a_n} & \vec{b_n} \end{bmatrix}$$

1.4.2 Example:

Suppose that $A = \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$ and $\vec{x} = \binom{2}{3}$

1. The following product can be written as a linear combination of vectors:

$$A\vec{x} = 2\begin{pmatrix}1\\0\end{pmatrix} + 3\begin{pmatrix}0\\-3\end{pmatrix} = \begin{pmatrix}2\\-9\end{pmatrix}$$

2. Is $\vec{b} = \binom{2}{9}$ in the span of the columns of A?
   If $\vec{b} \in Span\{A\}$, then $\vec{b} = c_1\binom{1}{0} + c_2\binom{0}{-3}$. This is true for $\vec{c} = \binom{2}{-3}$ so $\vec{b} \in Span\{A_{col}\}$

The equation $A\vec{x} = \vec{b}$has a solution iff $\vec{b}$ is a linear combination of the columns of A

1.4.3 Example: For what vectors $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ does the equation have a solution?

$$\begin{bmatrix} 1 & 3 & 4 \\ 2 & 8 & 4 \\ 0 & 1 & -2 \end{bmatrix} \vec{x} = \vec{b}$$

Solution:

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ 2 & 8 & 4 & b_2 \\ 0 & 1 & -2 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 2 & -4 & b_2 - 2b_1 \\ 0 & 1 & -2 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 2 & -4 & b_2 - 2b_1 \\ 0 & 0 & 0 & b_3 - \frac{1}{2}b_2 + b_1 \end{bmatrix}$$

So

$$\vec{b} = \begin{bmatrix} -\frac{1}{2}b_2 + b_3 \\ b_2 \\ b_3 \end{bmatrix}$$

Essential concept: If A is an $m \times n$ matrix, the following statements are logically equivalent – for a particular A, *all* are true or *all* are false:

- For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
- Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A
- The columns of A span $\mathbb{R}^m$
- A has a pivot position in every row

1.4.4 Summary: Ways of representing Linear Systems

1. A list of equations
2. An augmented matrix
3. A vector equation
4. A matrix equation

1.4.5 Matrix-vector products

If A is an $m \times n$ matrix, $\mathbf{u}$ and $\mathbf{b}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then:

- $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$
- $A(c\mathbf{u}) = c(A\mathbf{u})$

Moreover, the product $A\mathbf{x}$ can be easily calculated by taking advantage of the nature of a Matrix-vector product:

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_1x_1 + a_2x_2 + a_3x_3 \\ a_4x_1 + a_5x_2 + a_6x_3 \\ a_7x_1 + a_8x_2 + a_9x_3 \end{bmatrix}$$

The identity matrix: denoted $\mathbf{I}$, this $n \times n$ matrix contains 1's on the diagonal and 0's elsewhere, creating the universal property that $\mathbf{I_n x} = \mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$

1.5 TOPIC 5: Solution sets of Linear Systems

Homogenous: characteristic of linear systems of the form $A\vec{x} = \vec{b}$, $\vec{b} = \vec{0}$

Inhomogeneous: systems of the form $A\vec{x} = \vec{b}$, $\vec{b} \neq \vec{0}$

Because homogenous systems always have trivial solutions, the interesting question comes in asking whether they have non-trivial solutions.

$A\vec{x} = 0$ has a nontrivial solution $\iff$ there is a free variable $\iff$ A has a column with no pivot

1.5.1 Example: Identify the free variables and the solution set for

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 0 \\ 2x_1 - x_2 - 5x_3 &= 0 \\ x_1 - 2x_3 &= 0 \end{aligned}$$

$$\begin{bmatrix} 1 & 3 & 1 & 0 \\ 2 & -1 & -5 & 0 \\ 1 & 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & -7 & -7 & 0 \\ 0 & -3 & -3 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Row 3 has no pivot (x_3 is free) so there is a non-trivial solution.

$$\begin{cases} x_1 = & 2x_3 \\ x_2 = & -x_3 \\ x_3 & \text{is free} \end{cases} \implies x_3 \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

1.5.2 Parametric vector forms

Parametric vector form: a more convenient way of expressing the solutions of a linear system, taking advantage of the geometric interpretation of a linear system. In general, for free variables $x_k, ...x_n$ of $A\vec{x} = 0$, the solutions can all be written as

$$\vec{x} = \sum_{n=k}^{n} x_n \vec{v_n}$$

In other words, solving an equation amounts to finding an explicit description of the solution plane as a set spanned by $\mathbf{u}$ and $\mathbf{v}$. Thus, the equation from earlier describing the solution set can also be written as

$$\mathbf{x} = s\mathbf{u} + t\mathbf{v} \quad (s, t \in \mathbb{R})$$

emphasising the role of the free variables as arbitrary scalar multiples of the vectors forming a plane.

1.5.3 Example: Describe all solutions of $A\mathbf{x} = \mathbf{b}$

$$A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix} \text{ and } \quad \mathbf{b} = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus $x_1 = -1 + \frac{4}{3}x_3, x_2 = 2$, and x_3 is free. The general form vector is

$$\mathbf{x} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} \implies \mathbf{x} = \mathbf{p} + x_3\mathbf{v}$$

By replacing x_3 with a general parameter $t \in \mathbb{R}$, we have a universal solution for nonhomogenous systems:

$\mathbf{x} = \mathbf{p} + t\mathbf{v}$ is a solution for $Ax = b$ which is parallel to the line of the solution set $Ax = 0$ because it is a translation of $\mathbf{v}$ by the particular solution $\mathbf{p}$

1.5.4 Theorem:

Suppose the equation $Ax = b$ is consistent for some given $\mathbf{b}$ and let $\mathbf{p}$ be a solution. The solution set of $Ax = b$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$ where $\mathbf{v}_h$ is any solution of the homogenous equation $Ax = b$

1.6 TOPIC 6: Linear Independence

A set of vectors $\vec{v_1}, ..., \vec{v_k} \in \mathbb{R}^n$ are linearly independent if $\sum_{n=1}^{k} = c_k \vec{v_k} = 0$ has only the trivial solution ($\vec{c} = \vec{0}$). It is linearly dependent otherwise.

Establishing linear independence is thus equivalent to asking whether the equation $V\vec{c} = 0$ ($V \in \mathbb{R}^k :=$ the matrix corresponding to the linear combinations of the vectors) is only true for $\vec{c} = \vec{0}$

Example: For what values of h is the set of vectors linearly independent?

$$\begin{bmatrix} 1 \\ 1 \\ h \end{bmatrix}, \begin{bmatrix} 1 \\ h \\ 1 \end{bmatrix}, \begin{bmatrix} h \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & h & 0 \\ 1 & h & 1 & 0 \\ h & 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & h & 0 \\ 0 & h-1 & 1-h & 0 \\ 0 & 1-h & 1-h^2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & h & 0 \\ 0 & h-1 & 1-h & 0 \\ 0 & 0 & 2-h-h^2 & 0 \end{bmatrix}$$

If $2-h-h^2 = 0$ then we have a free variable and the vectors will be linearly dependent (because there will be more solutions than the trivial case).

Factoring, we get

$$0 = -(h+2)(h-1)$$

so for the vectors to be independent,

$$h \neq \{-2, 1\}$$

1.6.1 Linear Independence Theorems

1. More Vectors Than Elements For vectors $\vec{v}_1, ..., \vec{v}_k \in \mathbb{R}^n$, if $k > n$, then $\{\vec{v}_1, ..., \vec{v}_k\}$ is linearly dependent (because not every column of the matrix $A = (\vec{v}_1, ..., \vec{v}_k)$ would be pivotal).
2. Set Contains Zero Vector If any one or more of $\vec{v}_1, ..., \vec{v}_k$ is $\vec{0}$, then $\{\vec{v}_1, ..., \vec{v}_k\}$ is linearly dependent (again because there would be non-pivotal columns of the corresponding matrix).
3. A set containing only one vector is linearly independent iff $\mathbf{v}$ is not the zero vector
4. Sets of two vectors can be determined to be linearly dependent by inspection if one is a multiple of the other.

5. In geometric terms, two vectors are linearly dependent iff they lie on the same lie through the origin.

1.7 TOPIC 7: Introduction to Linear Transformations

Instead of as a linear combination of variables, matrices can also be viewed as a *function* from one set of vectors to another. Thus, solving the equation $Ax = b$ amounts to finding all vectors $\mathbf{x}$ in $\mathbb{R}^4$ that are transformed into the vector $\mathbf{b}$ in $\mathbb{R}^2$ under the "action" of multiplication by A.

Transformation (function, mapping): a rule that assigns to each vector $\mathbf{x} \in \mathbb{R}^n$ a vector $T(x) \in \mathbb{R}^m$.

Domain: The set $\mathbb{R}^n$ is the domain of T ($T \in \mathbb{R}^n$)

Codomain: The set $\mathbb{R}^m$ is the codomain of T ($T \in \mathbb{R}^n$)

This domain-codomain relationship is expressed in the notation $T : \mathbb{R}^n \to \mathbb{R}^m$

Image: this is the "resultant" vector of the transformation

$$T(x) \in \mathbb{R}^m \text{ for } \mathbf{x} \in \mathbb{R}^n \implies$$

"The vector T(x) is the image of x under the action of T"

Range: The set of all images of T(x) is the range of T. This is equivalent to the Span of the columns of a matrix.

For example, $f(x) = \sin x : \mathbb{R} \mapsto \mathbb{R}$, so the domain is $\mathbb{R}$, the codomain is $\mathbb{R}$, and the range is [-1, 1]

A function $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is linear if

- $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) \quad \forall \vec{u}, \vec{v} \in \mathbb{R}^n$
- $T(c\vec{v}) = cT(\vec{v}) \quad \forall \vec{v} \in \mathbb{R}^n, c \in \mathbb{R}$

Principle of superposition:

$$T(c_1\vec{v}_1 + ... + c_k\vec{v}_k) = c_1T(\vec{v}_1 + ... + c_kT(\vec{v}_k)$$

Every matrix transformation T_A is linear.

1.7.1 Matrix Transformations

A matrix transformation (a mapping where $\forall \mathbf{x} \in \mathbb{R}^n, T(x) = A\mathbf{x} \quad (A \in \mathbb{R}^{m\times n})$) can be written as $\mathbf{x} \mapsto A\mathbf{x}$
Similarly, every matrix transform is completely determined by what it does to the columns of the $n \times n$ identity matrix.

Geometric interpretations of transforms in $\mathbb{R}^2$:

1. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is a reflection through $x_1 = x_2$
2. $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is a projection onto the x-axis ($\begin{bmatrix} x_1 \\ 0 \end{bmatrix}$)
3. $A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$ is a scaling by k

Geometric interpretations of transforms in $\mathbb{R}^2$:

1. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is a projection onto the x_1, x_2-plane ($T(\vec{x}) = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix}$)
2. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is a reflection through the x_1, x_3-plane ($T(\vec{x}) = \begin{bmatrix} a \\ -b \\ c \end{bmatrix}$)

Standard vectors in $\mathbb{R}^n$:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Multiplying a matrix by $\vec{e}_i$ gives column i of A

1.7.2 Standard transform matrices

Counterclockwise rotation by angle θ about (0, 0):

$$T(x) = A\vec{x} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Reflection through x_1-axis:

$$T(x) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Reflection through x_2-axis:

$$T(x) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Reflection through $x_2 = x_1$:

$$T(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Reflection through $x_2 = -x_1$

$$T(x) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

Horizontal contraction:

$$T(x) = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}, \quad |k| < 1$$

Horizontal expansion:

$$T(x) = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}, \quad |k| > 1$$

Vertical contraction:

$$T(x) = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}, \quad |k| < 1$$

Vertical expansion:

$$T(x) = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}, \quad |k| > 1$$

Horizontal shear (left):

$$T(x) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, \quad k < 0$$

Horizontal shear (right):

$$T(x) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, \quad k > 0$$

Vertical shear (down):

$$T(x) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}, \quad k < 0$$

Vertical shear (up):

$$T(x) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}, k > 0$$

Projection onto the x_1-axis:

$$T(x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Projection onto the x_2-axis:

$$T(x) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Onto transform: a transform for which $A\vec{x} = \vec{b}$
- This is an existence property, for any $\vec{b} \in \mathbb{R}^m, A\vec{x} = \vec{b}$ has a solution
- T is onto iff its standard matrix has a pivot in every row

One-to-one transform: a transform for which there is at most one solution to $A\vec{x} = \vec{b}$
- This uniqueness property does not assert existence for all $\vec{b}$
- T is one-to-one iff the only solution to $T(\vec{x}) = \vec{0}$ is $\vec{x} = \vec{0}$
- T is one-to-one iff every column of A is pivotal

Onto means that the columns of A span $\mathbb{R}^n$

One-to-one means that the only solution is the trivial one and A has linearly independent columns

1.7.3 Determining the image of a transform

Example: Suppose T is a transform $T : \mathbb{R}^2 \mapsto \mathbb{R}^3$ such that

$$T(\vec{e}_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad T(\vec{e}_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

As any matrix can be written as a linear combination

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2$$

And since T is a linear transformation

$$\begin{aligned} T(\vec{x}) &= x_1 T(\vec{e}_1) + x_2 T(\vec{e}_2) \\ &= x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix} \end{aligned}$$

1.7.4 Example: Constructing a standard matrix

Define a linear transformation by

$$T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$$

$$A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) \end{bmatrix}$$

$$T(e_1) = T(1,0) = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix}$$

$$T(e_2) = T(0,1) = \begin{bmatrix} 1 \\ 7 \\ 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix}$$

Because there are more columns than rows, it cannot be onto. Because the two columns are not multiples of each other by inspection, the columns are linearly independent so the matrix is one-to-one.

2 Module 2: Matrix Algebra

2.1 TOPIC 1: Matrix Operations

Main diagonal: the diagonal entries $a_{11}, a_{22}, a_{33}, ...$ of $A = [a_{ij}], \quad A \in \mathbb{R}^{m \times n}$

Diagonal matrix: a square matrix $n \times n$ whose nondiagonal entries are zero (e.g. I_n)

Zero matrix: a matrix whose entries are all zero, written 0

Equal: property of two matrices of the same size with corresponding entries equal

Properties of Matrix addition:

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $r(A + B) = rA + rB$
5. $(r + s)A = rA + sB$
6. $r(sA) = (rs)A$

If a matrix B multiplies a vector $\mathbf{x}$, it is a transform $B\mathbf{x}$. If this vector is then multiplied by A, the composite transformed vector is $A(Bx)$. Thus:

$$A(Bx) = (AB)x$$

$$AB = A \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & ... & \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & ... & A\mathbf{b}_p \end{bmatrix}$$

Multiplication of matrices corresponds to composition of linear transformations.

Row column rule for matrix multiplication: If $A \in \mathbb{R}^{m \times n}$ has rows $\vec{a}_i$, and $B \in \mathbb{R}^{n \times p}$ has columns $\vec{b}_j$, each element of the product $C = AB$ is the dot product $c_{ij} = \vec{a}_i \cdot \vec{b}_j$

Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

The number of columns of A must match the number of rows in B in order for a linear combination such as $A\mathbf{b}_1$ to be defined.

Properties of Matrix multiplication:

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(B + C)A = BA + CA$
4. $r(AB) = (rA)B = A(rB)$
5. $I_m A = A = AI_n$
6. If $AB = AC$ it is usually NOT true that $B = C$
7. If AB is the zero matrix, you CANNOT conclude that either $A =$ or $B = 0$

If $A \neq 0$ and $\mathbf{x} \in \mathbb{R}^n$, then $A^k\mathbf{x}$ is the result of left-multiplying $\mathbf{x}$ by A repeatedly k times. $(A^0 = I_n)$

Transpose: the transpose of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose columns are the corresponding rows of A.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Transposition properties:

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(rA)^T = rA^T$
4. $(AB)^T = B^T A^T$

Hence, *the transpose of a product of matrices equals the product of their transposes in the reverse order.*

2.2 TOPIC 2: The Inverse of a Matrix

Invertible: a square matrix is invertible if there is a matrix $C \in \mathbb{R}^{n\times n}$ such that

$$CA = I \quad \text{and} \quad AC = I$$

Here, C is the unique inverse of A often denoted A^-1

Singular matrix: a noninvertible matrix

Nonsingular matrix: an invertible matrix

2.2.1 Finding the inverse

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Determinant: $\det A = ad - bc$. If $\det A = 0$, A is not invertible.

If A is an invertible $n \times n$ matrix, then for each $\vec{b}$ in $\mathbb{R}^n$, the equation $Ax = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$

While the equation above can used to solve an equation $Ax = b$, row reduction of $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ is almost always faster.

Properties of invertible matrices:

1. $(A^-1) - 1 = A$
2. $(AB)^-1 = B^-1A^-1$
3. $(A^T)^-1 = (A^-1)^T$

Elementary matrix: a matrix that is obtained by performing a single elementary row operation on an identity matrix.

If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

Every elementary matrix is invertible and square where the inverse of E is the elementary matrix that transforms E back into I.

An $n \times n$ matrix is invertible iff A is row equivalent to I_n, in which case any sequence of elementary row operations which maps $A \mapsto I_n$ also transforms $I_n \mapsto A^{-1}$

A has an inverse iff for all $\vec{b} \in \mathbb{R}^n$, $Ax = b$ has a unique solution

2.2.2 To find A^{-1}:

Row reduce the augmented matrix $\begin{bmatrix} A & I \end{bmatrix}$ If A is row equivalent to I, then $\begin{bmatrix} A & I \end{bmatrix}$ is row equivalent to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$

Example: Find the inverse of $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$

Solution:

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -\frac{9}{2} & 7 & -\frac{3}{2} \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{bmatrix}$$

Since $A \sim I$, A is invertible and

$$A^{-1} = \begin{bmatrix} -\frac{9}{2} & 7 & -\frac{3}{2} \\ -2 & 4 & -1 \\ \frac{3}{2} & -2 & \frac{1}{2} \end{bmatrix}$$

Row reduction of $\begin{bmatrix} A & I \end{bmatrix}$ to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$ can be viewed as the simultaneous solution of the n systems

$$A\mathbf{x} = \mathbf{e_1}, \quad A\mathbf{x} = \mathbf{e_2}, \quad ..., \quad A\mathbf{x} = \mathbf{e_n}$$

2.3 TOPIC 3: Characterisations of Invertible Matrices

Invertible matrix theorem: for a given square matrix A, the following are all true or all false:

1. A is an invertible matrix
2. A is row equivalent to I_n
3. A has n pivot positions
4. The equation $Ax = 0$ has only the trivial solution
5. THe columns of A form a linearly independent set
6. The linear transformation $x \mapsto Ax$ is one-to-one
7. The equation $Ax = b$ has at least one solution for each $\mathbf{b} \in \mathbb{R}^n$
8. The columns of A span $\mathbb{R}^n$
9. The linear transformation $x \mapsto Ax$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$
10. There is an $n \times n$ matrix C such that $CA = I$
11. There is an $n \times n$ matrix D such that $AD = I$
12. A^T is an invertible matrix

If A and B are two square matrices and $AB = I$, then both A and B are invertible with $B = A^{-1}$ and B^{-1}

Invertible linear transformation: a transform is invertible if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S(T(x)) = x \quad \forall x \in \mathbb{R}^n \tag{1}$$
$$T(S(x)) = x \quad \forall x \in \mathbb{R}^n \tag{2}$$
$$\tag{3}$$

If such S exists, it is unique and must be a linear transform. This transform is invertible iff its standard matrix A is nonsingular. In that case, $S(x) = A^{-1}x$ is the unique function satisfying (1) and (2).

Ill-conditioned matrix: an invertible matrix that can become singular if some of its entries are changed just slightly. This becomes dangerous if roundoff error makes a nonsingular matrix appear invertible.

2.4 TOPIC 4: Partitioned Matrices

Partitioned matrix: a matrix whose entries are meant to be considered differently than a list of column vectors based on the system it represents. For example, the matrix

$$A = \left[\begin{array}{ccc|cc|c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -2 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

can also be the 2×3 partitioned matrix

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

whose entries are the blocks (submatrices)

$$A_{11} = \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, \qquad A_{12} = \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, \qquad A_{13} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$
$$A_{21} = \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, \qquad A_{22} = \begin{bmatrix} 1 & 7 \end{bmatrix}, \qquad A_{23} = \begin{bmatrix} -4 \end{bmatrix}$$

Assuming the blocks of two partitioned matrices are the same size, addition between them is defined as the block-by-block sum of corresponding partitions. Similarly, multiplication by a scalar is performed piecewise.

Partitioned matrices can be multiplied by the usual row-column rule as if the block entires were scalars, provided that for a product AB, the column partition of A matches the row partition of B.

Conformable: property of two matrices that are partitioned the same way

Block diagonal matrix: a partitioned matrix with zero blocks off the main diagonals - such a matrix is invertible iff each block on the diagonal is invertible

Partitioned matrices are especially useful in numerical solutions of especially large systems where partitioning can make more efficient use of system resources

2.5 TOPIC 5: MATRIX FACTORISATIONS

Factorization: an equation that expresses a matrix A as a product of two or more matrices.

Whereas multiplication was a synthesis of data, factorisation is an analysis.

2.5.1 LU Factorisation

This is a more efficient process for solving the system

$$Ax = b_1, \quad Ax = b_2, \quad ..., \quad Ax = b_p$$

The method:

1. Assume A is an $m \times n$ matrix that can be row reduced to echelon form without row interchanges
2. A can be written in the form $A = LU$ where L is an $m \times n$ lower triangular matrix with 1's on the diagonal and U is an $m \times n$ echelon form of A

$$A = LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix} \begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3. When $A = LU, Ax = b$ can be written as $L(Ux) = b$
4. We can then find x by solving the pair of equations

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

5. First solve for y, then for x. Because L and U are triangular, both equations are easy to solve.

Example: Given

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & - & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 1 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

use this LU factorisation to solve $Ax = b$, where $b = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$

Solution:

$$\begin{bmatrix} L & b \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 1 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} I & y \end{bmatrix}$$

Then,

$$\begin{bmatrix} U & y \end{bmatrix} = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & -1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

In this method, finding x requires 28 arithmetic operations. Meanwhile, row reduction of $[A \quad b]$ to $[I \quad x]$ takes 62 operations.

2.5.2 An LU factorisation algorithm

Suppose A can be reduced to an echelon form U using only row replacements that add a multiple of one row to another row *below it*. In this case, there exist unit lower triangular elementary matrices such that

$$E_p...E_1A = U$$

Then

$$A = (E_p...E_1)^{-1}U = LU$$

where

$$L = (E_p...E_1)^{-1}$$

It can be shown that the row operations which reduce A to U also reduce L to I, because $E_p...E_1L = (E_p...E_1)(E_p...E_1)^{-1} = I$

In summary:

1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
2. Place entries in L such that the same sequence of row operations reduces L to I.

Example: Find an LU factorization of

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$$

Solution: Since A has four rows, L should be 4×4. The first column of L is the first column of A divided by the top pivot entry

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

Next row reduce A to its echelon form U:

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U$$

The columns eliminated in each step of the row reduction process divided by the pivot form L:

$$\begin{bmatrix} 2 & & & \\ -4 & 3 & & \\ 2 & -9 & 2 & \\ -6 & 12 & 4 & 5 \end{bmatrix} \Longrightarrow L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}$$

2.6 TOPIC 6: The Leontief Input-Output Model

The interactions between sectors can be represented by the matrix equation

$$(I - C)x = d$$

Where: I is the identity matrix C is a coefficient matrix representing the demand of each sector for goods from each other sector $\mathbf{x}$ is a unit consumption vector describing the number of units of output which are routed to each sector based on the consumption matrix $\mathbf{d}$ is the final demand vector (the output of the system after accounting for intermediate demand, Cx)

If C and $\mathbf{d}$ have nonnegative entries and if each column sum of C is less than 1,

$$\mathbf{x} = (I - C)^{-1}\mathbf{d}$$

and

$$(I - C)^{-1} \approxeq I + C + C^2 + ...C^m$$

which, in most real-world circumstances, will have powers of the consumption matrix approach 0 quite quickly.

2.7 TOPIC 7: Computer Graphics

Often computer graphics will be represented by a list of points which can be connected by straight lines.

2.7.1 2D Graphics

A particular graphic, for example the letter "N" can be stored as coordinates in a data Matrix

$$D = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 0.5 & 0.5 & 6 & 6 & 5.5 & 5.5 & 0 \\ 0 & 0 & 6.42 & 0 & 8 & 8 & 1.58 & 8 \end{bmatrix}$$

This representation makes the transformation or animation of pictures very easy: the resultant figure is described by the matrix product AD where A is any linear transformation.

Unfortunately, translation is not a linear transform so cannot be represented by matrix multiplication.

Homogeneous coordinates: corresponding sets of coordinates where $\forall x, y \in \mathbb{R}^2, (x, y) \mapsto (x, y, 1) \in \mathbb{R}^3$ For example, (0, 0) has homogenous coordinates (0, 0, 1). These coordinates are not multiplied or added by scalars but can be transformed via multiplication by 3×3 matrices.

With homogeneous coordinates we are able to represent translations:

$$(x, y) \mapsto (x + h, y + k) \implies (x, y, 1) \mapsto (x + h, y + k, 1)$$

Thus,

$$\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x+h \\ y+k \\ 1 \end{bmatrix}$$

More broadly, any linear transformation in $\mathbb{R}^2$ can be represented in homogeneous coordinates by a partitioned matrix of the form $\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}$ where A is a 2×2 standard matrix.

2.7.2 3D Graphics

Homogeneous coordinates for the point (x, y, z) can usually be expressed as (X, Y, Z, H) where

$$x = \frac{X}{H}, \quad y = \frac{Y}{H}, \quad z = \frac{Z}{H}$$

2.8 TOPIC 8: Subspaces of $\mathbb{R}^n$

Subspace: any set H in $\mathbb{R}^n$ that has the properties:

- The zero vector is in H
- $\forall \mathbf{u}, \mathbf{v} \in H \quad \mathbf{u} + \mathbf{v} \in H$
- $\forall \mathbf{u} \in H, \quad c\mathbf{u} \in H$

A subspace is closed under addition and scalar multiplication.

One of the most common ways of visualising a subspace is a plane through the origin.

Example: If $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ and $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, then H is a subspace of $\mathbb{R}^n$

Column space: the subspace of a matrix A is the set Col A of all linear combinations of the columns of A

Col A equals $\mathbb{R}^m$ only when the columns of A span $\mathbb{R}^m$. Otherwise, Col A is only part of $\mathbb{R}^m$

A vector b is in Col A iff the equation Ax = b has a solution.

Null space: the null space of a matrix A is the set Nul A of all solutions of the homogeneous equation Ax = 0

The null space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$. Equivalently, the set of all solutions of a system $Ax = 0$ of m homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.

Basis: for a subspace H of $\mathbb{R}^n$, the basis is a linearly independent set in H that spans H

Standard basis: the set $\{\mathbf{e}_1, ..., \mathbf{e}_2\}$

The pivot columns of a matrix A form a basis for the column space of A. Note: the columns of an echelon form B are often not in the column space of A

2.9 TOPIC 9: Dimension and Rank

The main reason for a selecting a basis for a subspace H rather than just a spanning set is that each vector in H can bbe written in only one way as a linear combination of the basis vectors.

Coordinate vector: for each x in a subspace H,

$$[x]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

for $c_1, ...c_p$ such that

$$x = c_1\mathbf{b}_1 + ... + c_p\mathbf{b}_p$$

Isomorphism: a correspondence $x \mapsto [x]_{\mathfrak{B}}$ which is a one-to-one mapping which preserves linear combinations.

Dimension: the dimensions of a nonzero subspace H, denoted dim H, is the number of vectors in any basis for H

The space $\mathbb{R}^n$ has dimension n so every basis for $\mathbb{R}^n$ consists of n vectors. Similarly, a plane through **0** is two-dimensional while a line through **0** is one-dimensional.

Rank: the rank of a matrix A, denoted rank A, is the dimension of the column space of A

Since the pivot columns of A form a basis for Col A, the rank of A is just the number of pivot columns in A.

The Rank Theorem: If a matrix A has n columns, then rank A + dim Nul A = n

The Basis Theorem: Any linearly independent set of exactly p elements in a p-dimensional subspace of $\mathbb{R}^n$ H is automatically a basis for H.

2.9.1 Extension of the Invertible Matrix Theorem

For an $n \times n$ matrix A, the following statements are equivalent to the statement that A is nonsingular:

1. The columns of A form a basis of $\mathbb{R}^n$
2. Col A $= \mathbb{R}^n$
3. dim Col A = n
4. rank A = n
5. Nul A $= \{\mathbf{0}\}$
6. dim Nul A = 0

3 Module 3: Determinants, Eigenvalues, and Markov Chains

3.1 TOPIC 1: INTRODUCTION TO DETERMINANTS

A 2×2 matrix is invertible iff its determinant is nonzero.

For a 1×1: $\det A = a_{11}$

For a 2×2: $\det A = a_{11}a_{22} - a_{12}a_{21}$

For a 3×3:

$$\Delta = a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

This is equivalent to

$$\Delta = a_1 1 \cdot \det A_{11} - a_{12} \cdot \det A_{12} + a_{13} \cdot \det A_{13}$$

Where A_{11}, A_{12}, A_{13} are obtained from A by deleting the first row and one of the three columns. (And where A_{ij} is the matrix formed by deleting the i-th row and j-th column of A).

In general, an $n \times n$ determinant is defined by the determinants of $(n-1) \times (n-1)$ submatrices.

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$$

3.1.1 Example:

Compute the determinant of

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Solution:

$$\begin{aligned} \det A &= 1 \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5 \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0 \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} \\ &= 1(0-2) - 5(0) + 0(-4-4) \\ &= -2 \end{aligned}$$

(i, j)-cofactor of A: the number given by $C_{ij} = (-1)^{i+j} \det A_{ij}$

Cofactor expansion across the i-th row of A: a method of computing the determinant with use of cofactors

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + ... + a_{in}C_{in}$$

The plus or minus in the (i, j)-cofactor depends on the position of a_{ij} in the matrix:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

3.1.2 Example:

Use a cofactor expansion across the third row to compute $\det A$ for

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Solution:

$$\begin{aligned} \det A &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \\ &= a_{31}(-1)^{3+1} \det A_{31} + a_{32}(-1)^{3+2} \det A_{32} + a_{33}(-1)^{3+3} \det A_{33} \\ &= 0 + (-2)(-1) \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0 \\ &= 2(-1) = -2 \end{aligned}$$

The cofactor approach is particularly useful in matrices with many zeros because those cofactors need not be calculated.

If A is a triangular matrix, then det A is the product of the entries on the main diagonal of A

Numerical note: in general, a cofactor expansion requires more than $n!$ multiplications

3.2 TOPIC 2: Properties of Determinants

For a square matrix A:

1. Row replacement operations on A to produce a matrix B make $\det B = \det A$
2. Row interchange operations on A to produce B make $\det B = -\det A$
3. Row scaling operations make $\det B = k \det A$

Using these operations, we can reduce a matrix to echelon form and use the fact that the determinant of a triangular matrix is the product of its diagonal entries to more efficiently calculate the determinant.

3.2.1 Example:

Compute det A for

$$A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$$

Solution:

$$\det A = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix}$$

As replacement does not change the determinant and interchange only changes the sign,

$$\det A = -(1)(3)(-5) = 15$$

If a square matrix A has been reduced to echelon form U by row replacements and r interchanges alone,

$$\det A = (-1)^r \det U$$

Further,

$$\det A = \begin{cases} (-1)^r \cdot \prod_{i=1}^{n} U_{ii} & \text{when A is invertible} \\ 0 & \text{when A is not invertible} \end{cases}$$

A square matrix A is invertible iff $\det A \neq 0$

Numerical note: the row reduction method of calculating the determinant takes only about $2n^3/3$ operations compared to the $n!$ of cofactor expansion.

If A is an $n \times n$ matrix, then $\det A^T = \det A$

3.2.2 Determinants and Matrix Products

Multiplicative property: For square matrices A and B,

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

NOTE: $\det(A + B)$ is usually *not* equal to $\det A + \det B$

3.3 TOPIC 3: Cramer's rule, volume, and linear transformations

Cramer's rule: Let A be an invertible $n \times n$ matrix and $A_i(\vec{b})$ be the matrix obtained from A by replacing column i by the vector $\vec{b}$. For any $\vec{b}$ in $\mathbb{R}^n$, the unique solution of $A\vec{x} = \vec{b}$ has entries given by

$$x_i = \frac{\det A_i(\vec{b})}{\det A}, \quad i = 1, 2, ..., n$$

Example: Use Cramer's rule to solve the system

$$3x_1 - 2x_2 = 6 \tag{4}$$
$$-5x_1 + 4x_2 = 8 \tag{5}$$

Solution:

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, \quad A_1(\vec{b}) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, \quad A_2(\vec{b}) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

Since $\det A = 2$ the system has a unique solution. By Cramer's rule:

$$x_1 = \frac{\det A_i(\vec{b})}{\det A} = \frac{24 + 16}{2} = 20$$
$$x_2 = \frac{\det A_i(\vec{b})}{\det A} = \frac{24 + 30}{2} = 27$$

Laplace transforms: an approach, useful in many engineering problems, which involves converting an appropriate system of linear differential equations into a system of linear algebraic equations whose coefficients involve a parameter.

Adjugate (classical adjoint) of A: the transpose of the matrix of cofactors of A

$$\text{adj}A = \begin{bmatrix} C_{11} & C_{21} & ...C_{n1} & \\ C_{12} & C_{22} & ...C_{n2} & \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & ... & C_{nn} \end{bmatrix}$$

For an invertible square matrix A,

$$A^{-1} = \frac{1}{\det A}\text{adj}A$$

Cramer's rule is a theoretical tool for studying the tolerance of the solution to Ax = b to changes in entry b or A. It is sometimes useful for hand calculation of 3×3 matrices with complex entries. Usually, *it is hopelessly inefficient.*

3.3.1 Determinants as Area or Volume

If A is 2×2, the area of the parellelogram determined by the columns of A is $|\det A|$.

If A is 3×3, the volume of the parralelepiped determined by the columns of A is also $|\det A|$

Example: Calculate the area of the parellelogram determined by the points (-2, -2), (0, 3), (4, -1), and (6, 4)

Solution: First, translate the parellelogram to the origin by subtracting (-2, -2) from each of the other vertices. The new vertices have coordinates (0, 0), (2, 5), (6, 1), and (8,6). This is in the Span of

$$A = \begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}$$

And $|\det A| = -28$, so the area of the parellelogram is 28.

3.3.2 Linear transformations

Let $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$. If S is a parallelogram in $\mathbb{R}^2$ and $T(S)$ is the set of images of points of S,

$$\{areaofT(S)\} = |\det A|\{areaofS\}$$

This statement is analogous for volume in $\mathbb{R}^3$

3.4 TOPIC 4: Markov Chains

Probability vector: a vector with nonnegative entries that add up to 1

Stochastic matrix: a square matrix whose columns are probability vectors

Markov Chain: a model used to describe an experiment or measurement which is performed many times in the same way, where the outcome of each trial depends only on the immediately preceding trial It can also be represented as a sequence of probability vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2...$ together with a stochastic matrix P such that

$$\mathbf{x}_1 = P\mathbf{x}_0, \quad \mathbf{x}_2 = P\mathbf{x}_1, \quad \mathbf{x}_3 = P\mathbf{x}_2, ...$$

Thus the Markov chain is described by the first-order difference equation

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{for } k = 0, 1, 2, ...$$

In general, $\mathbf{x}_k = P^k\mathbf{x}_0$

When a Markov chain of vectors is in $\mathbb{R}^n$, the entries in $\mathbf{x}_k$ (called the *state vector*) list the probabilities that the system is in each of n possible states or that the outcome of the experiment is one of n possible outcomes.

Steady-state vector (equilibrium vector): If P is a stochastic matrix, then a steady-state vector for P is a probability vector $\mathbf{q}$ such that

$$P\mathbf{q} = \mathbf{q}$$

Every stochastic matrix has a steady-state vector

To find the steady-state vector solve the equation $P\mathbf{x} = \mathbf{x}$:

$$\begin{aligned} P\mathbf{x} - \mathbf{x} &= 0 \\ P\mathbf{x} - I\mathbf{x} &= 0 \\ (P - I)\mathbf{x} &= 0 \end{aligned}$$

Which can be solved by row reduction to find the general solution.

Next, choose a simple basis for the solution space and divide that by the sum of its entries to find a probability vector q in the set of solutions of $Px = x$.

Regular: a quality of a stochastic matrix where some matrix power of P^k contains only strictly positive entries.

If P is an $n \times n$ regular stochastic matrix, then P has a unique steady-state vector q. Further, if x_0 is any initial state and $x_{k+1} = Px_k$ for k = 0, 1, 2, ... then the Markov chain $\{x_k\}$ converges to q as $k \to \infty$

3.5 TOPIC 5: Eigenvalues and Eigenvectors

Eigenvector: a nonzero vector x such that $Ax = \lambda x$ for some $n \times n$ matrix A and scalar λ

Eigenvalue: a scalar λ for which there is a nontrivial solution of $Ax = \lambda x$ (where x is an eigenvector corresponding to λ)

Example: Show that 7 is an eigenvalue of $\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$

7 is an eigenvalue of A iff

$$Ax = 7x$$

Thus,

$$\begin{aligned} Ax - 7x &= 0 \\ (A - 7I)x &= 0 \\ A - 7I &= \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix} \end{aligned}$$

The columns of A are linearly dependent so 7 *is* an eigenvalue of A. The eigenvectors are:

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

So, every vector of the form $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $(x_2 \neq 0)$ is an eigenvector corresponding to $\lambda = 7$

■

Warning: row reduction can be used to find eigenvectors but not eigenvalues. An echelon form of matrix A usually does not display the eigenvalues of A.

λ is an eigenvalue of an $n \times n$ matrix A iff the equation $(A - \lambda I)x = 0$ has a nontrivial solution.

The set of all solutions of that equation is just the nullspace of $A - \lambda I$ so this set is a subspace of $\mathbb{R}^n$

Eigenspace: the subspace of $\mathbb{R}^n$ consisting of the zero vector and all the eigenvectors corresponding to λ

A scalar λ will be an eigenvalue of A if $(A - \lambda I)x = 0$ has free variables

3.5.1 Difference equations:

Eigenvectors can be used to construct solutions of the first-order difference equation

$$x_{k+1} = Ax_k \quad (k = 0, 1, 2, ...)$$

A solution of this equation is an explicit description o the sequence $\{x_k\}$ whose formula for each x_k does not depend directly on A or on the preceding terms in the sequence. The simplest way to build this solution is to take an eigenvector x_0 and a corresponding eigenvalue such that

$$x_k = \lambda^k x_0$$

Which is a solution because

$$Ax_k = A(\lambda^k x_0) = \lambda^k (Ax_0) = \lambda^k (\lambda x_0) = \lambda^{k+1} x_0 = x_{k+1}$$

Eigenvalue Theorems:

1. The eigenvalues of a triangular matrix are the entries on its main diagonal
2. 0 is an eigenvalue of A $\iff$ A is not invertible
3. Stochastic matrices have an eigenvalue equal to 1
4. If $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ are eigenvectors that correspond to distinct eigenvalues, then those eigenvectors are linearly independent

3.6 TOPIC 6: The Characteristic Equation

A scalar λ is an eigenvalue of an $n \times n$ matrix A iff λ satisfies the characteristic equation

$$\det(A - \lambda I) = 0$$

(Because a matrix is invertible iff 0 is not an eigenvalue of a square matrix A and the determinant is also not 0 by the Invertible Matrix Theorem)

Characteristic polynomial of A: the expanded form of $\det(A - \lambda I)$ for an $n \times n$ matrix which will be a polynomial of degree n

Trace: the sum of the diagonal elements of its matrices

When a (2×2) matrix is singular, one eigenvalue is 0 and the other is the trace.

Warnings:

1. Matrices can have the same eigenvalues but not be similar
2. Similarity is not the same as row equivalence
3. Row operations on a matrix usually change its eigenvalues

Algebraic multiplicity: the multiplicity of an eigenvalue as a root of the characteristic polynomial

Geometric multiplicity: the dimension of Null$(A - \lambda I)$

- Geometric multiplicity is always at least 1

For an $n \times n$ matrix A with a_i as the algebraic multiplicity of λ_i and g_i the geometric multiplicity:

1. $1 \leq a_i \leq n$
2. $1 \leq g_i \leq a_i$

3.6.1 Similarity

If A and B are square matrices, then A is similar to B if there is a matrix P such that $P^{-1}AP = B$. If so, A and B have the same characteristic polynomial and hence the same eigenvalues with the same multiplicities.

This can lead to a more efficient algorithm for calculating large powers of A^k

3.7 TOPIC 3: Diagonalization