Systems of Linear Equations

- A system of equations of the form:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots\ldots\ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots\ldots\ldots\ldots\ldots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad\qquad\qquad\qquad \vdots \qquad \vdots &\quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots\ldots\ldots\ldots + a_{mn}x_n &= b_m
\end{aligned}
$$

where the elements a_{ij} and b_i are scalars and the x_j are "unknown" variables is called **a system of m linear equations in n unknowns.**

- Any (ordered) n-tuple $(s_1, s_2, \ldots\ldots, s_n)$ of scalars which satisfies all of the equations is called a **solution** of the system. The set of all solutions is called the **solution set** of the system.

An Example of a Linear System

- Consider the system:

$$\begin{aligned} 5x_2 + 10x_3 + 8x_4 &= 23 \\ x_1 + 2x_2 + 6x_3 + 7x_4 &= 16 \\ 2x_1 + 4x_2 + 12x_3 + 6x_4 &= 24 \end{aligned}$$

- **It is a system of 3 linear equations in 4 unknowns.**

Matrix Formulation

- A system of linear equations can be more compactly expressed in matrix notation as:

  $A\mathbf{x} = \mathbf{b}$, where $\mathbf{A} = [a_{ij}]$ is called the coefficient matrix, and

  $$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$ are vectors.

- Recall that a vector is an ordered k-tuple of scalars. Vectors are notated in various ways: $(x_1, x_2, \ldots, x_k)$
  or $[x_1\ x_2\ \ldots\ x_k]$ (*referred to as a row vector*)

  $$\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}$$ *(referred to as a column vector)*

Example System in Matrix Form

- The earlier example system can be expressed as the system A**x** = **b**, where

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

and **b** = $\begin{bmatrix} 23 \\ 16 \\ 24 \end{bmatrix}$

Vector Formulation

- A system of linear equations can also be expressed in a vector form:

  $x_1\mathbf{v_1} + x_2\mathbf{v_2} + \ldots\ldots. + x_n\mathbf{v_n} = \mathbf{b}$, where the x_i are scalar unknowns and the $\mathbf{v_i}$ are column vectors formed from the coefficients of the original linear system.

- This formulation can be interpreted as: if we can find scalars x_i satisfying the equation, then the given vector **b** can be expressed in terms of the given vectors $\mathbf{v_i}$. This formulation is not useful for solving the system, but will become very important when we are working with vectors.

Example System in Vector Form

- The earlier example system can be expressed as the system $x_1\mathbf{v_1} + x_2\mathbf{v_2} + x_3\mathbf{v_3} + x_4\mathbf{v_4} = \mathbf{b}$, where

$$\mathbf{v_1} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}, \mathbf{v_3} = \begin{bmatrix} 10 \\ 6 \\ 12 \end{bmatrix}, \mathbf{v_4} = \begin{bmatrix} 8 \\ 7 \\ 6 \end{bmatrix}$$

and $\mathbf{b} = \begin{bmatrix} 23 \\ 16 \\ 24 \end{bmatrix}$

SUMMARY

- A linear system can be expressed in **three** different forms:
- ***Equation Form***: a set of m linear equations in n unknowns
- ***Matrix Form***: $A\mathbf{x} = \mathbf{b}$, where A is an m×n matrix (coefficient matrix), **x** is an n-vector of unknowns and **b** is an m-vector
- ***Vector Form***: $x_1\mathbf{v_1} + x_2\mathbf{v_2} + \ldots\ldots + x_n\mathbf{v_n} = \mathbf{b}$, where the x_i are scalar unknowns and the $\mathbf{v_i}$ are column m-vectors of (fixed) coefficients.
- In practice, linear systems can arise in any of the three formulations. However, we will prefer to use the matrix formulation most of the time.

SOME NOTATION - 1

- Some important sets:
  - $\mathbb{N}$ = natural numbers = {0,1,2,3, ….. }
  - $\mathbb{Z}$ = integers
  - $\mathbb{Z}^+$ = positive integers
  - $\mathbb{R}$ = real numbers ($\mathbb{R}^+$ = positive real numbers)
  - $\mathbb{Q}$ = rational numbers ($\mathbb{Q}^+$ = positive rational numbers)
  - $\mathbb{C}$ = complex numbers
  - For any set X, $\mathbb{P}(X)$ = power set of X = 2^X
- **Note**: The font used for these special sets in my ppt's will always be Castellar. We will try to avoid using these letters for other sets or objects.

SOME NOTATION - 2

- In keeping with the above, we will use $\mathbb{R}^n$ to stand for the cartesian product of $\mathbb{R}$ taken with itself n times ($n \geq 1$). $\mathbb{R}^n$ is also, therefore, the set of all n-vectors with real entries (sometimes also called coefficients or coordinates). $\mathbb{R}^n$ will often be referred to as **euclidean** n-space. Occasionally we will work with vectors with complex entries; the notation for the set of all n-vectors with complex entries is $\mathbb{C}^n$.

SOME NOTATION - 3

- Most of the time we will be working with linear systems in which the coefficients are real numbers, i.e. systems of the form $A\mathbf{x} = \mathbf{b}$, where A is an $m \times n$ matrix with real entries, and $\mathbf{b}$ is an m-vector with real entries. Any solution is an n-vector with real entries. In terms of the above notation, the solution set of the system is a subset of $\mathbb{R}^n$.

Solving Linear Systems

- Small systems of linear equations (with two or three variables) can be solved by a method of "elimination" or a method of "substitution". We wish to present a more systematic strategy which can be used in a *mechanical* way to deal with any system. However, we are going to do this in a somewhat roundabout way. We will work directly with matrices, and develop a fundamental **matrix algorithm, which has several different applications.** Solving linear systems is just one of the uses of this algorithm.
- **Remark**: In the process of solving a linear system, the variables play no real role. All calculations are done with the coefficient matrix and the RHS scalars. So it makes sense that a purely matrix algorithm can do the task for us.

Elementary Row Operations

- Given any $m \times n$ matrix A, we define three **elementary row operations:**
  - Multiplication of one row of A by a **non-zero** scalar c (**scale**)
  - Replacement of one row of A by the sum of the row and a scalar multiple of a **different** row (**replace**)
  - Interchange of two rows of A (**interchange**)
- **Observation 1:** So by applying an elementary row operation e to A, we get a **new** matrix e(A).
- **Observation 2:** To each elementary row operation e, there corresponds an elementary row operation e_1 of the same type such that $e_1(e(A)) = A$. In other words, the process is reversible.

2 Special Types of Matrices - 1

- An m×n matrix is said to be in **echelon form** if:
  - All non-zero rows are above any all-zero rows
  - Each leading entry (***i.e. first non-zero entry***) of a row is to the right of the leading entry of the row above it
  - All entries in a column below a leading entry are zero

- *NB: Actually, the third condition above follows from the second. However, we have written it out explicitly here in the interest of clarity.*

2 Special Types of Matrices - 2

- An m×n matrix is said to be a **reduced row** echelon matrix or **in row-reduced echelon form** (RREF) if:
  - All non-zero rows are above all zero rows
  - Each leading entry (*i.e. the first non-zero entry*) of a row is to the right of the leading entry of the row above it
  - The leading entry (*note again: first non-zero entry*) in each non-zero row is 1
  - Each column which contains such a leading entry (necessarily 1) has *all its other entries as 0*
- In other words, an RREF matrix is in echelon form and has two further requirements also.