MTH 264 Introduction to Matrix Algebra - Summer 2023.

LN2A. Matrix Equations and Row Operations.

These lecture notes are mostly lifted from the text **Matrix and Power Series, Lee and Scarborough, custom 5th edition**. This document highlights parts of the text that are used in the lecture sessions.

Often, we run into the problem of decomposing a vector $\mathbf{v}$ into a linear combination of some set of vectors $A = \{\mathbf{a_1}, ..., \mathbf{a_n}\}$. This problem can be translated in terms of matrix equations.

Theorem 2A.1. Vector Decomposition in terms of Matrix Equations

Let $\mathbf{v} \in \mathbb{R}^m$ and let $A = \{\mathbf{a_1}, ..., \mathbf{a_n}\} \subset \mathbb{R}^m$ be a finite set of vectors. Then, the problem of finding all scalars $k_1, ..., k_n \in \mathbb{R}$ such that $\mathbf{v} = k_1\mathbf{a_1} + ... + k_n\mathbf{a_n}$ is equivalent to solving the matrix equation

$$\mathbf{Ak} = \mathbf{v} \qquad \text{or equivalently,} \qquad k_1\mathbf{a_1} + \cdots + k_n\mathbf{a_n} = \begin{pmatrix} | & | & & | \\ \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} \\ | & | & & | \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = \mathbf{v}$$

with $\mathbf{v}$ represented as a column vector, $(\mathbf{a_1}, ..., \mathbf{a_n})$ representing the columns of $\mathbf{A} \in \mathbb{R}^{m\times n}$, and $\mathbf{k} \in \mathbb{R}^n$ with entries $(k_1, ..., k_n)$.

One way to justify this result is to consider factoring the set of scalars from the linear combination of vectors. Here, we understand factoring as the reversal of matrix multiplication. We can easily confirm that that intuition is true by doing the matrix multiplication operation.

One application of this involves systems of linear equations.

Theorem 2A.2. Systems of Linear Equations correspond to Matrix Equations

A linear system with m equations and unknowns $x_1, ..., x_n$ corresponds to a matrix equation $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. This correspondence is given by:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + ... + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + ... + a_{2,n}x_n = b_2 \\ \qquad\qquad\qquad\qquad \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + ... + a_{m,n}x_n = b_m \end{cases} \iff \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

We also provide another result that characterizes the cardinality/number of solutions to a matrix equation of this type. Before that, we introduce some notation that the author forgot to include in the previous lecture notes.

Definition 2A.3. Span of a Set of Vectors

Let $V = \{\mathbf{v_1}, ..., \mathbf{v_k}\} \in \mathbb{R}^n$ be some set of vectors. Then, the **span** of V, denoted span(V), is the set of all linear combinations of the vectors $\mathbf{v_1}, ..., \mathbf{v_k}$. That is,

$$\text{span}(V) = \{a_1\mathbf{v_1} + ... + a_k\mathbf{v_k} \text{ such that } a_1, ..., a_k \in \mathbb{R}\}$$

We also define the span of the empty set to be the set with only the zero vector. That is, $\text{span}(\emptyset) = \{\mathbf{0}\}$.

With this, we can characterize the solution sets of matrix equations.

Theorem 2A.4. Cardinality of Solutions to $\mathbf{Ax} = \mathbf{b}$

Let $\mathbf{Ax} = \mathbf{b}$ be a matrix equation. Then, the set of solutions $\mathbf{x}$ that satisfy this equation follow exactly one of the following results:

(a) The matrix equation has **no solutions**.

(b) The matrix equation has **exactly one solution**.

(c) The matrix equation has **infinitely many solutions** with the solution set described by a span of some set of vectors $\mathbf{v_1}, ..., \mathbf{v_k}$ translated by some vector $\mathbf{b}$. The set of vectors is generally not unique.

For this course, we discuss two ways to solve an equation in the form $\mathbf{Ax} = \mathbf{b}$. The first method is by using inverses.

Theorem 2A.5. Solving Matrix Equations by Inverses

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ and let $\mathbf{b} \in \mathbb{R}^n$ be given such that $\mathbf{A}$ is invertible with inverse $\mathbf{A}^{-1}$. Then, the equation $\mathbf{Ax} = \mathbf{b}$ has exactly one solution given by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Observe that this method only applies if two conditions are satisfied: $\mathbf{A}$ is a square matrix and $\mathbf{A}$ is invertible. These conditions are very restrictive and this method only applies to a very small set of matrix equations. Fortunately, there is another result that would help us solve a bigger family of matrix equations.

Theorem 2A.6. Equivalence of Solutions

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and let $\mathbf{b} \in \mathbb{R}^m$ be given. Let $\mathbf{B} \in \mathbb{R}^{m\times m}$ be some invertible matrix. Then, $\mathbf{x} \in \mathbb{R}^n$ is a solution to $\mathbf{Ax} = \mathbf{b}$ if and only if $\mathbf{x}$ is a solution to $\mathbf{BAx} = \mathbf{Bb}$. In other words, the solutions to $\mathbf{Ax} = \mathbf{b}$ are exactly the solutions to $\mathbf{BAx} = \mathbf{Bb}$.

In other resources, the theorem above is stated as the following: If $\mathbf{A}$ and $\mathbf{M}$ are similar matrices, the solutions of the equation $\mathbf{Ax} = \mathbf{b}$ are exactly the solutions to $\mathbf{Mx} = \mathbf{b}$. Here, we define that matrices $\mathbf{A}$ and $\mathbf{M}$ are **similar** if and only if there exists an invertible matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{M}$.

This result is very helpful since, with a special set of invertible matrices, we can solve matrix equations in the form $\mathbf{Ax} = \mathbf{b}$ algorithmically. For this course, we won't explicitly state what these invertible matrices are since we're not really interested in finding $\mathbf{B}$. Instead, we simulate multiplication by $\mathbf{B}$ using **row operations** – which we'll define after introducing more results.

Definition 2A.7. Augmented Matrix

The corresponding **augmented matrix** of the matrix equation $\mathbf{Ax} = \mathbf{b}$ is the matrix given by

$$\begin{pmatrix} \mathbf{A} & \mathbf{b} \end{pmatrix} = \begin{pmatrix} | & & | & | \\ \mathbf{a_1} & \cdots & \mathbf{a_k} & \mathbf{b} \\ | & & | & | \end{pmatrix}$$

where the column vectors $\mathbf{a_1}, ..., \mathbf{a_k}$ describe the columns of $\mathbf{A}$.

The following result allows us to simulate left matrix multiplications on the augmented matrix instead of on the matrix equation. This will simplify the calculation and presentation of our solutions.

Theorem 2A.8. Simulation on Augmented Matrices

Let $\mathbf{M}$ be the augmented matrix for the equation $\mathbf{Ax} = \mathbf{b}$. Let $\mathbf{B}$ be some matrix such that $\mathbf{BA}$ is defined. Then, the matrix $\mathbf{BM}$ is the augmented matrix for the equation $\mathbf{BAx} = \mathbf{Bb}$

This result basically tells us that we can work on the level of augmented matrices instead of on the level of matrix equations. This also tells us that we can go back and forth between representations as necessary/convenient.

Finally, we introduce **row operations** which we'll simulate over the augmented matrix.

Theorem 2A.9. Row Operations

Let $\mathbf{Ax} = \mathbf{b}$ be a matrix equation. Let $\mathbf{M} = \begin{pmatrix} -\mathbf{R_1}- \\ \vdots \\ -\mathbf{R_m}- \end{pmatrix}$ be the corresponding augmented matrix described using row vectors $\mathbf{R_1}, ..., \mathbf{R_n}$. Then, the following **row operations** on $\mathbf{M}$ are equivalent to multiplication of the equation by some invertible matrix and therefore, preserves the solutions $\mathbf{x}$ of the equation.

- **Type (I)** Swap rows $\mathbf{R_i}$ and $\mathbf{R_j}$. We sometimes denote this operation using $\mathbf{R_i} \leftrightarrow \mathbf{R_j}$.
- **Type (II)** Replace row $\mathbf{R_i}$ with $k\mathbf{R_i}$ with some nonzero scalar $k \in \mathbb{R}$. We sometimes denote this operation using $k\mathbf{R_i} \mapsto \mathbf{R_i}$.
- **Type (III)** Replace row $\mathbf{R_i}$ with a linear combination $k_1\mathbf{R_1} + ... + k_n\mathbf{R_m}$ with $k_i = 1$, i.e. $\mathbf{R_i}$ must be in the linear combination. We sometimes denote this operation using $k_1\mathbf{R_1} + ... + k_n\mathbf{R_m} \mapsto \mathbf{R_i}$.

For this course, to solve matrix equations in a somewhat structured way, we introduce an algorithm called **Gaussian Elimination**. However, we must first define matrix forms.

Definition 2A.10. Pivots and Matrix Forms

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. Describe the rows of $\mathbf{A}$ using row vectors $\mathbf{R_1}, ..., \mathbf{R_m} \in \mathbb{R}^n$.

- **(a)** Define the **pivot** of the row $\mathbf{R_i}$, denoted as $\mathsf{Pivot}(\mathbf{R_i})$, as the first nonzero entry of $\mathbf{R_i}$ from left to right (i.e. column indices in increasing order).
- **(b)** Define the **pivot index** of the row $\mathbf{R_i}$, denoted as $\mathsf{PivotIndex}\,\mathbf{R_i}$, as the column index of $\mathsf{Pivot}\,\mathbf{R_i}$.
- **(c)** We say that $\mathbf{A}$ is in **row echelon form** if and only if (1) all zero rows are at the bottom of the matrix $\mathbf{A}$ and (2) the pivot indices are strictly increasing as we look at rows from top to bottom, i.e. $\mathsf{PivotIndex}\,\mathbf{R_1} > \mathsf{PivotIndex}\,\mathbf{R_2} > ... > \mathsf{PivotIndex}\,\mathbf{R_k}$ with $\mathbf{R_k}$ the last row that is not a zero row (i.e. we only consider rows for which $\mathsf{Pivot}\,\mathbf{R_i}$ is defined).
- **(d)** We say that $\mathbf{A}$ is in **reduced row echelon form** if and only if (1) $\mathbf{A}$ is in row echelon form; (2) all pivots are equal to 1; and (3) for columns containing a pivot, all entries except the pivot must equal 0.

For this course, we typically stop when the augmented matrix $\mathbf{M}$ is in **row echelon form** and then do a method called **back-substitution** which we'll define later.