The 18.C06 Master Document

Jack D.V. Carson

`jdcarson@mit.edu`

Contents

1 Linear Algebra

This document assumes you already have a reasonable familiarity with the most basic ideas of linear algebra (i.e. What is a matrix? What is matrix multiplication? Why would one be compelled to multiply a matrix?). From hence, we shall begin with the beginning.

1.1 The Basics

- In linear algebra, we most often want to express some situation in terms of a formula $Ax = b$ for a matrix A, and vectors x, b. Then, once we have expressed them in this form, we wish to solve for the values of x that give b. The simplest way to do this is with **Gaussian Elimination** wherein rows are added to one another or multiplied by scalar constants to give a **diagonal matrix**. This algorithm always goes columnwise. I.e. start from the first column and try and clear everything such that everything below the pivot is 0. This may not always be the fastest approach but it is consistent. From this diagonal matrix, it is simple to **backsubstitute** to get the complete values of x.

$$\begin{bmatrix} 1 & 3 & 1 & 9 \\ 1 & 1 & -1 & 1 \\ 3 & 11 & 6 & 35 \end{bmatrix} \xrightarrow[r_3-3r_1]{r_2-r_1} \begin{bmatrix} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 2 & 3 & 8 \end{bmatrix} \xrightarrow{r_3+r_2} \begin{bmatrix} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

  Here we have gone from the augmented matrix $[A|b]$ to a much nicer $[U|b]$. By recalling that matricies are just an abstraction of linear functions, we can look at the last row and say that $x_3 = 0$. We can *propogate* this upwards to say therefore that $-2x_2 + 0 = -8 \rightarrow x_2 = 4$

- If we get a 0 in a pivot position, we say that a matrix is **singular**. In this case $Ax = b$ may not have a solution, or it may have infinite solutions. But it certainly does not have a single solution. All linear equations of this type either have 0, 1, or ∞ solutions.

- It wouldn't hurt us to brush up on matrix multiplication either. It turns out there are many ways to think about this operation.

  1. **Entry-wise:** For each $1 \le i \le m$ and $1 \le jp$ we have

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

2. **Inner product:** C_{ij} is the dot product (also known as "inner product") of the ith row in A and the jth column in B. For instance,

$$C = \begin{bmatrix} - & x_1 & - \\ - & x_1 & - \\ & \vdots & \\ - & x_1 & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} x_1 \cdot b_1 & x_1 \cdot b_2 & \cdots & x_1 \cdot b_p \\ x_2 \cdot b_1 & x_2 \cdot b_2 & \cdots & b_2 \cdot b_p \\ \vdots & \vdots & \ddots & \vdots \\ x_m \cdot b_1 & x_m \cdot b_2 & \cdots & x_m \cdot b_p \end{bmatrix}$$

3. **Column-wise:** the jth column of a matrix C is the matrix-vector product of A and the jth column of B. For instance,

$$C = A \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ Ab_1 & Ab_2 & \cdots & Ab_p \\ | & | & & | \end{bmatrix}$$

4. **Outer product:** C is the sum of the product of ith column of A and the ith row of B such as

$$C = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} - & y_1 & - \\ - & y_2 & - \\ & \vdots & \\ - & y_n & - \end{bmatrix} = \sum_{i=1}^{n} \begin{bmatrix} | \\ a_i \\ | \end{bmatrix} \begin{bmatrix} - & y_i & - \end{bmatrix}$$

Properties of Matrix Multiplication

1. **Associative:** $A(BC) = (AB)C$

2. **Distributive:** $A(B + C) = AB + AC \iff (A + B)C = AC + BC$

3. **Non-commutative:** $AB \neq BA$

4. **Identity:** $IA = AI = A$

- Although it can often be lost in the abstraction of mathematics, matricies really, truly are **linear operators**. They transform spaces and vectors to other spaces and other vectors. As an example consider 2023 Recitation 1 Problem 1

Example: 1.1

1. Find a 2×2 matrix such that when you multiply a 2-D vector by it, the result is a reflection of the vector across the origin

2. Find a 3×3 matrix such that when you multiply a 3-D vector by it, it swaps the second and third coordinates.

3. If you have a 4×4 matrix A, find a 4-D vector x such that Ax is the second column of A.

Solution (1.1).

1. To reflect across the origin, $x \mapsto -x$, $y \mapsto -y$. Therefore, $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$.

2. Here $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x \\ z \\ y \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + z \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

   Here A is a **permutation matrix**, as it permutes one or more of the variables. Permutation matricies have a variety of desirable properties such as $A^{-1} = A^T$

3. Since we are only interested in the second column, we want the products $\mathbf{c_i} x_i = 0$ for $i = 1, 3, 4$ to be 0. The only way to guaruntee this is to specify $x_{1,3,4} = 0$ and $x_2 = 1$. Therefore $\mathbf{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^T$

Takeaways (1.1).

1. All of the matrix formulations here are useful to recognize, particularly the intuition of part (3). This will be seen many times.

2. One useful way to interpret the result in part (2) is the fact that left-multiplied matricies operate on the rows of that which they multiply. Right-multiplied matricies operate on columns. Because $\mathbf{x}$ has only one column, right-multiplication is nonsensical.

Example: 1.2

1. Find a 3×3 matrix P such that in $B = PA$ is the result of subtracting the second row from the third row of A and then swapping the first and second rows.

2. Find a 4×4 matrix that right multiplies A such that result $C = AQ$ is A after dividing the first column by two, and then adding the first column to the second and third columns

3. Does the order of performing the operations in (1) and (2) matter?

Solution (1.2).

1. $P = \underbrace{\begin{bmatrix} 1 & & \\ & 1 & \\ & -1 & 1 \end{bmatrix}}_{r_3 = r_3 - r_2} \underbrace{\begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix}}_{\text{swap } \mathbf{c_1}, \mathbf{c_2}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}$

2. $Q = \begin{bmatrix} 1/2 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. This multiplication should be obvious by the column-wise definition above.

3. Although it might look like we are dealing with the commutative propety here, we are really dealing with the associative property. In (1). $B = (P_1 P_2)A = P_1(P_2 A)$. So the order of operations does not matter. In (2) it is the same $C = A(Q_1 Q_2) = (AQ_1)Q_2$

Takeaways (1.2).

– Remember that left multiplication always affects rows only, and that right multiplication affects columns only. Complex operations can be formed by chaining linear operators together.

– Matrix multiplication is associative! When doing these chained operations, the order does not matter.

– Remember the forms that these kinds of matricies take. They are not always obvious.

- Before we close up, there are some other noteworthy operations we can perform with matricies that will follow us around. The **transpose** of a matrix A^T "flips" a matrix A such that

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

and with it we can define certain identities such as $\boxed{(A^T)_{i,j} = A_{j,i}}$ and $\boxed{(AB)^T = B^T A^T}$. Finally, we can define the dot product between two vectors $\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b}$. All three of these facts will show up constantly.

- Last but not least, the **inverse** A^{-1} of a matrix is the matrix such that $AA^{-1} = A^{-1}A = I$. More interestingly we can say that in a situation $Ax = b$, $x = A^{-1}b$. This matrix inverse does not always exists, and is exceedingly impractical to calculate for large matricies, an issue we will deal with thoroughly in optimization. A matrix is said to be **invertible** if and only if it is square and full column rank. That is, every column has a pivot. The inverse, if it exists, is always unique. Like the transpose it is subject to the identity $\boxed{(AB)^{-1} = B^{-1}A^{-1}}$

- We can define more facts that will help us with these two operations. For instance $(A^T)^T$ and $(A + B)^T = A^T + B^T$, which follow directly from the definition of the transpose. More useful is $\boxed{(A^{-1})^T = (A^T)^{-1}}$. All of these identities will help us greatly.

1.2   Vector Spaces

Definition

A **vector space** V is a set of elements (e.g., vectors in ^{n}, polynomials, diagonal 2×2 matricies) defined over a "field" F of scalars that are closed under

1. **Vector Addition:** For any vectors $\mathbf{u}$ and $\mathbf{v}$ in V, $\mathbf{u} \pm \mathbf{v}$ also belongs the vector

---

> space V.
>
> 2. **Scalar Multiplication:** For any vector $\boldsymbol{v}$ in V and scalar c in F, $c\boldsymbol{v}$ is an element of the vector space

- If $W \subseteq V$ is also a vector space with respect to the operations in V, W is a **subspace** of V. Specifically for any $v, w \in W$, $v + w \in W$ and $cv \in W$. For instance of we define a vector space ^{2}. W could be ^{2} or $W = \{0\}$. A more enlightening sample vector space is a line that passes through the origin (note that a line that does not pass through the origin would not be a vector space since $\alpha v \notin V$ for $\alpha = 0$). Every point on this line is closed under addition and scalar multiplication.

- There are an infinite number of obscure facts to say about spaces and subspaces. More interestingly, is how to show something is a space or a substapce. We can be convinced that for subspaces S_1 and S_2 $S = S_1 \cap S_2$ is also a subspace by showing that S is closed under addition and scalar multiplication.

  Proof. Assuming that this is true, any $v, w \in S$ will also be $v, w \in S_1$ and $v, w \in S_2$ by the definition of an intersection. Then $v + w$ will be in S_1 and S_2 by the deifnition of subspace. Thus $v + w \in S_1 \cap S_2 = S$ which shows that it is closed under addition. Then we can define $v \in S_1 \cap S_2$ and $\alpha \in$. Since S_1 is a subspace, $\alpha v \in S_1$ and S_2 such that $\alpha v \in S_1 \cap S_2 = S$, showing that it is closed under scalar multiplication.

Column Space

- The **column space** of an $m \times n$ matrix A, denoted $C(A)$ is the set of linear combinations of columns of A, also known as the **span** of A. We can define useful facts such as

  1. Formally, $C(A) = \{Ax \mid x \in^n\}$
  2. $Ax = b$ has a solution if and only if $b \in C(A)$
  3. If $m = n$, then A is invertible if and only if $C(A) =^n$