Math 2270-1

Notes of 9/30/2019

Review

- Recall: Exam 2 on Wednesday. Will cover chapters 2 and 3.
- Tomorrow, more review, driven by your questions.

The following list is neither self contained nor is it complete. Rather, the individual items should stir your memory about facts, concepts, and connections. If some do not then it is a good idea to review the associated material!

Chapter 2 Review

- We add two $m \times n$ matrices entry by entry, just like we add vectors.
- Similarly, we multiply any matrix by a scalar by multiplying each entry with that scalar.

However, we multiply matrices so that the product of two matrices is the standard matrix of the composition of the corresponding linear functions.

Matrix Multiplication

- The composition $f \circ g$ of two linear functions f and g is linear, and its matrix is the product of the matrices of the constituent functions.

$$
\begin{array}{c}
\overbrace{\mathbb{R}^p \ \underset{B}{\overset{g}{\longrightarrow}} \ \mathbb{R}^n \ \underset{A}{\overset{f}{\longrightarrow}} \ \mathbb{R}^m}^{f \circ g} \\
\quad n \times p \qquad\quad m \times n \\
C = AB \\
C \text{ is } m \times p
\end{array}
$$

- Note the switch in the sequence. B comes first in the diagram and second in the product, just like g comes first in the diagram and second in the composition.

Six Views of $C = AB$

- We'll look at six different ways of thinking about matrix multiplication. All of them are useful!

- For any matrix A let $\mathbf{r}_i(A)$ denote the i-th row of A, interpreted as a matrix with one row, and let $\mathbf{c}_i(A)$ denote the i-th column, interpreted as matrix with one column. We also identify 1×1 matrices with their single scalar entry.

- Let's illustrate the descriptions with the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 4 & 5 \\ 10 & 11 \end{bmatrix}.$$

1. **The Formula**. Here is what you might find in a textbook or mathematical dictionary: The product of an $m \times n$ matrix A and an $n \times p$ matrix B is an $m \times p$ matrix $C = AB$ where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, p.$$

- For our example,

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \times 2 + 2 \times 1 & 1 \times 1 + 2 \times 2 \\ 2 \times 3 + 1 \times 4 & 1 \times 3 + 2 \times 4 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 5 \\ 10 & 11 \end{bmatrix}. \end{aligned}$$

2. **Writing it**. We saw in class that it is advantageous to write the second factor to the upper right of the first factor. The product fits into the corner made by the two factors, and the $i-j$ entry of C sits at the intersection of the i-th row of A and the j-th column of

B. In our example we get

$$\begin{array}{cc} & \begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix} \\ \begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix} & \begin{bmatrix}4 & 5\\ 10 & 11\end{bmatrix}\end{array}$$

- More generally, we get:

$$B \quad n \times p$$

$$\begin{bmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{i1} & \dots & b_{ij} & \dots & b_{ip} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nj} & \dots & b_{np} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix} \qquad \begin{bmatrix} & & & & \\ & & \vdots & & \\ & \dots & c_{ij} & \dots & \\ & & \vdots & & \\ & & & & \end{bmatrix}$$

$$A \quad m \times n \qquad\qquad C = AB \quad m \times p$$

It is evident from this picture that

- the $i - j$ entry of C is the (dot) product of the i-th row of A and the j-th column of B,
- the j-th column of C is the product of A and the j-th column of B,

– the i-th row of C is the product of the i-th row of A and B.

- Following is an elaboration of these views. But first, here is a clean copy of the same picture:

$$
\begin{array}{cc}
 & B \quad n \times p \\
 & \begin{bmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{i1} & \dots & b_{ij} & \dots & b_{ip} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nj} & \dots & b_{np} \end{bmatrix} \\
\begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix} & \begin{bmatrix} & & & & \\ & & \vdots & & \\ & \dots & c_{ij} & \dots & \\ & & \vdots & & \\ & & & & \end{bmatrix} \\
A \quad m \times n & C = AB \quad m \times p
\end{array}
$$

3. The entry by entry view.

$$c_{ij} = \mathbf{r}_i(A)\mathbf{c}_j(B).$$

In our example

$$\begin{aligned}
AB &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} & \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} 4 & 5 \\ 10 & 11 \end{bmatrix}.
\end{aligned}$$

4. The Column View. This is actually how we first derived our formula for matrix multiplication: the j-th column of C equals A multiplied with the j-column of B. As a formula:

$$\mathbf{c}_j(C) = A\mathbf{c}_j(B), \qquad j = 1, \dots, p.$$

In our example:

$$\begin{aligned}
C &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} A \begin{bmatrix} 2 \\ 1 \end{bmatrix} & A \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} 4 & 5 \\ 10 & 11 \end{bmatrix}.
\end{aligned}$$

- Note that in this view every column of the product is a linear combination of the columns of A. The coefficients of the linear combination are in the corresponding column of B.

5. **The Row View**. The i-th row of C is the i-th row of A multiplied with B:

$$r_i(C) = r_i(A)B.$$

In our example:

$$\begin{aligned} C &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} B \\ \begin{bmatrix} 3 & 4 \end{bmatrix} B \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \\ \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 4 & 5 \\ 10 & 11 \end{bmatrix} \end{aligned}$$

- Note that in this view every row of the product is a linear combination of the rows of B. The coefficients of the linear combination are in the corresponding row of A.

6. **The matrix view**. Note that the product of the k-th column of A and the k-th row of B is an $m \times p$ matrix, the product of an $m \times 1$ matrix and a $1 \times p$ matrix. The $i - j$ entry of $\mathbf{c}_k(A)\mathbf{r}_k(B)$ is $a_{ik}b_{kj}$. So we get, by our formula

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj},$$

that

$$C = \sum_{k=1}^{n} \mathbf{c}_k(A)\mathbf{r}_k(B).$$

In our example

$$\begin{aligned} C &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 5 \\ 10 & 11 \end{bmatrix} \end{aligned}$$

in general, the product of an $m \times 1$ matrix A and a $1 \times p$ matrix B is an $m \times p$ matrix C. Every row of C is a multiple of A and every column of C is a multiple of B.

The Inverse of a Matrix

- A square $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that
$$CA = I \quad \text{and} \quad AC = I$$
where $I = I_n$, the $n \times n$ identity matrix.

- If A is not invertible it is said to be **singular**.

- the terms invertible and singular only apply to square matrices.

- If A is invertible its inverse is unique.

- There are many equivalent properties of a square matrix that are equivalent to invertibility. Some of them are expressed in what the textbook calls the

Invertibility Theorem Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.

b. A is row equivalent to the identity matrix.

c. A has n pivot positions.

d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

e. The columns of A form a linearly independent set.

f. The linear transformation $\mathbf{x} \longrightarrow A\mathbf{x}$ is one-to-one.

g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

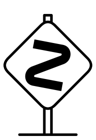
h. The columns of A span $\mathbb{R}^n$

i. The linear transformation maps $\mathbb{R}^b$ onto $\mathbb{R}^n$.

j. There is an $n \times n$ matrix C such that $CA = I$.

k. There is an $n \times n$ matrix D such that $CA = I$. (Of course, the left and right inverses C and D are actually equal.)

l. A^T is an invertible matrix.



We will soon extend this list.

More Properties of Inverse Matrices

- If A is invertible and $A\mathbf{x} = \mathbf{b}$ then $\mathbf{x} = A^{-1}\mathbf{b}$.
- If A is invertible and $AB = C$ then $B = A^{-1}C$.
- If A is invertible and $BA = C$ then $B = CA^{-1}$.
- The process of inverting and transposing a matrix commute:

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T.$$

- Assuming A and B are invertible and have the same size,

$$(AB)^{-1} = B^{-1}A^{-1}$$

  In general

$$B^{-1}A^{-1} \neq A^{-1}B^{-1}$$

  since matrix multiplication does not commute.

Partitioned Matrices

- A **partitioned matrix**, or **block matrix**, is a matrix whose entries are matrices. Block matrices have many useful properties but we discussed only one

Major Principle: Provided the partitions are conformable, multiplying block matrices works exactly like multiplying matrices.

- As an application, we computed the inverse of a 2×2 block matrix:

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

  where A_{11} is $p \times p$, A_{22} is $q \times q$, and $p + q = n$. Of course, A_{12} is $p \times q$ and A_{21} is $q \times p$.

Matrix Factorizations

- A **factorization** of a matrix A is obtained by writing A as a product of several (usually 2 or 3) matrices.
- In particular we discussed the LU factorization of a square matrix

$$A = LU \tag{1}$$

where

L is **unit lower triangular**, i.e.,

$$l_{ii} = 1, \quad i = 1, 2, \dots n \quad \text{and} \quad j < i \quad \Longrightarrow \quad a_{ij} = 0$$

U is **upper triangular**, i.e.,

$$l_{ii} = 1, \quad i = 1, 2, \dots n \quad \text{and} \quad j > i \quad \Longrightarrow \quad a_{ij} = 0$$

- In other words, denoting possibly non-zero entries by x, L and U are of the form:

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ x & 1 & 0 & \dots & 0 & 0 \\ x & x & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x & x & x & \dots & 1 & 0 \\ x & x & x & \dots & x & 1 \end{bmatrix}$$

and

$$U = \begin{bmatrix} x & x & x & \dots & x & x \\ 0 & x & x & \dots & x & x \\ 0 & 0 & x & \dots & x & x \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & x & x \\ 0 & 0 & 0 & \dots & 0 & x \end{bmatrix}$$

- A major application of the LU factorization is to solve the linear system $A\mathbf{x} = \mathbf{b}$ by solving the two triangular systems

$$L\mathbf{y} = \mathbf{b} \quad \text{and} \quad U\mathbf{x} = y.$$

Computing the LU factorization is equivalent to applying row operations to convert A to upper triangular form.

- Frequently we need to interchange rows during Gaussian Elimination. The process is called **pivoting**. Pivoting can be expressed in terms of a **pivot matrix** which is a matrix that has been obtained from the identity matrix by permuting its rows or columns.

It usually is a bad idea actually to compute an inverse matrix.

Chapter 3: Determinants

- Determinants are numbers associated with **square** matrices. Throughout these notes let A be an $n \times n$ square matrix.
- We denote the determinant of A by $|A|$ or $\det A$. Note that in this context the vertical bars do not mean absolute values.
- The determinant of a 1×1 matrix equals its unique entry.
- For a 2×2 matrix we define
$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$
- For $n > 2$ the determinant of A is defined recursively. Suppose A is an $n \times n$ matrix where $n > 2$. We define A_{ij} to be the $(n-1)\times(n-1)$ matrix obtained from A by removing the i-th row and the j-th column.
- For $n \geq 2$, the **determinant** of an $n \times n$ matrix $A = [\, a_{ij} \,]$ is
$$\begin{aligned} \det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + \ldots + (-1)^{1+n} a_{1n} \det A_{1n} \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j} \end{aligned} \quad . \tag{2}$$

However, a corresponding expansion gives the same numerical value for **any row or column**.

- The corresponding formula is usually expressed in terms of **cofactors**:

$$C_{ij} = (-1)^{i+j} \det A_{ij}.$$

- The factor

$$(-1)^{i+j}$$

creates the familiar checkerboard pattern

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

- **Theorem 1** on page on page 168 in the textbook gives the general formulas

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij} \tag{3}$$

for any choice of i and

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij} \tag{4}$$

for any choice of j.

- There is no proof of this fact in the textbook. One way to see this, with a bit of work, is to show that formula (2) implies that

$$\det A = \sum_{\sigma} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma_i} \tag{5}$$

where the sum goes over all $n!$ permutations of the set $\{1, 2, \ldots, n\}$ and the large symbol π indicates the product of n factors, one from each row i, and the column σ_i.

- The formula (5) is symmetric in the rows and columns, and so any row or column can be used to compute the determinant.

- Clearly, in the cofactor expansion (3) or (4) you want to pick rows or columns that contain many zeros.

- An extreme case of the exploitation of zero entries is provided by triangular matrices. **The determinant of a triangular matrix is the product of the diagonal entries.** (This is **Theorem 2** on page 169 of the textbook.)

- The formula (5) gives the determinant as the sum of $n!$ terms, each of which has n factors. Computing the determinant by that formula is prohibitively expensive even for small values on n.

- Determinants can be computed much more efficiently by row operations. The relevant facts are contained in Theorem 3 on page 171 of the textbook:

Theorem 3. (Row Operations) Let A be a square matrix. Then

a. If a multiple of a row of A is added to another row to produce a matrix B, then

$$\det B = \det A.$$

b. If two rows of A are interchanged to produce B then

$$\det B = -\det A.$$

c. If one row of A is multiplied by a scalar k to produce B then

$$\det B = k \det A.$$

- All of these statements can be proved by observing that they are true for $n = 2$ and then using induction based on the cofactor expansion.

Theorem 4, page 173.

s. A square matrix A is invertible if and only if

$$\det A \neq 0.$$

- This can be seen by reducing the matrix to row echelon form and applying Theorem 2 on triangular matrices.
- (The Label **s.** indicates the position in the invertible matrix theorem.)

Theorem 5, page 174. If A is a square matrix then

$$\det A = \det A^T.$$

- This is obvious by the fact that any row or column can be used for the computation of the determinant.

Theorem 6, page 175. If A and B are both $n \times n$ matrices then

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

- We saw that this is true by observing that the statement holds for elementary matrices (those implementing row operations) and writing one of the matrices as a product of elementary matrices and the identity matrix.

Linearity: The determinant is a linear function of each row or column separately.

- This follows immediately from the cofactor expansion.
- **Cramer's Rule** rule states that for the linear system
$$A\mathbf{x} = \mathbf{b}$$
the i-th entry of $\mathbf{x}$ is given by
$$x_i = \frac{|A_i(\mathbf{b})|}{|A|}$$

where $A_i(\mathbf{b})$ is the matrix formed by replacing the i-th column of A with $\mathbf{b}$.

- To see that this is true let I_i be the matrix obtained from the identity matrix by replacing the i-th column with $\mathbf{x}$. Thus

$$I_i = [\,\mathbf{e}_1 \quad \mathbf{e}_2 \quad \ldots \mathbf{e}_{i-1} \quad \mathbf{x} \quad \mathbf{e}_{i+1} \quad \ldots \quad \mathbf{e}_n\,]$$

Then, by the way we defined matrix multiplication,

$$AI_i = A_i.$$

The determinant of the product equals the product of the determinants:

$$|A||I_i| = |A_i|, \quad \text{i.e.,} \quad |I_i| = \frac{|A_i|}{|A|}.$$

Moreover, by expanding about the i-th row we see that

$$|I_i| = x_i.$$

Cramer's rule follows.

- In particular, Cramer's Rule gives a formula for A^{-1}

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \ldots & C_{n1} \\ C_{12} & C_{22} & \ldots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \ldots & C_{nn} \end{bmatrix}$$

- The matrix

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

is called the **adjugate** of A.

note that the adjugate is the **transpose** of the matrix of cofactors!

- The determinant can be interpreted geometrically as the volume V of a parallelotope defined by the columns of A:

$$V = |\det A|$$

where the vertical bars in this case do denote the absolute values.