Theorem 11. Let $A, A' \in \mathcal{M}_{m\times p}, B, B' \in \mathcal{M}_{p\times n}$. Then

(3) $A(B + B') = AB + AB'$

Proof: We choose some $i, j \in \mathbb{N}$ such that $i \leq m$ and $j \leq n$. Now we know from the definition of multiplication that

$$\begin{aligned}\mathrm{ent}_{ij}(A(B + B')) &= \sum_{k=1}^{p} a_{ik}(b_{kj} + b'_{kj}) \\ &= \sum_{k=1}^{p} (a_{ik}b_{kj}) + (a_{ik}b'_{kj}) \\ &= \sum_{k=1}^{p} a_{ik}b_{kj} + \sum_{k=1}^{p} a_{ik}b'_{kj}\end{aligned}$$

Again, the definition of multiplication leads us to

$$\mathrm{ent}_{ij}(A(B + B')) = \mathrm{ent}_{ij}(AB) + \mathrm{ent}_{ij}(AB')$$

Of course our choice of i and j were arbitrary, and so this is true for any i, j and thus $A(B+B') = AB + AB'$. $\square$

(4) $(cA)B = cAB = A(cB)$ for all $c \in \mathbb{R}$

Proof: As above we choose an arbitrary element from row i and column j of $(cA)B$. From the definitions of scalar and matrix multiplication, we have

$$\mathrm{ent}_{ij}((cA)B) = \sum_{k=1}^{p} (ca_{ik})b_{kj}$$

Using the commutative, distributive, and associative properties of the real numbers it is not hard to see that

$$\begin{aligned}\mathrm{ent}_{ij}((cA)B) &= c\sum_{k=1}^{p} a_{ik}b_{kj} \\ &= \sum_{k=1}^{p} a_{ik}(cb_{kj})\end{aligned}$$

But this means that $\mathrm{ent}_{ij}((cA)B) = \mathrm{ent}_{ij}(cAB) = \mathrm{ent}_{ij}(A(cB))$ for all i and j. Thus we have $(cA)B = cAB = A(cB)$ as required. $\square$

2.1.5 a. If A is an $m \times n$ matrix and $A\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$, show that $A = O$.

We proceed with a proof of the contrapositive. Let us start by assuming that $A \neq O$ and go on to show that in this case we can find some $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{c}$ with $\mathbf{c} \neq \mathbf{0}$.

Now if $A \neq O$ then there exists at least one entry a_{ij} such that $a_{ij} \neq 0$. We choose $\mathbf{x}$ to be equal to the row of A which contains a_{ij}. That is to say $\mathbf{x} = (a_{i1}, a_{i2}, \ldots, a_{in})$. Now let us examine $\mathbf{c}$. We know that $c_i = \sum\limits_{k=1}^{n} {a_{ik}}^2$. Of course we can't get a negative number by squaring a real number so ${a_{ik}}^2 \geq 0$. We also know that ${a_{ij}}^2$ in particular is strictly greater than zero. Just to be painfully clear, if we add a number that is at least zero to a number that is more than zero, then we will get a number that is more than zero. Thus we have $c_i > 0$ and so $\mathbf{c} \neq \mathbf{0}$. $\square$

b. If A and B are $m \times n$ matrices and $A\mathbf{x} = B\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$, show that $A = B$.

We may rewrite the equation as $A\mathbf{x} - B\mathbf{x} = \mathbf{0}$. This leads to $(A - B)\mathbf{x} = \mathbf{0}$. But from part (a) we know that if this is true for all $\mathbf{x}$ then we have $A - B = 0$ or $A = B$. $\square$

2.1.6 Prove or give a counterexample. Assume all the matrices are $n \times n$.

a. If $AB = CB$ and $B \neq O$ then $A = C$.
$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = O = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$

b. If $A^2 = A$ then $A = O$ or $A = I$.
$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

c. $(A + B)(A - B) = A^2 - B^2$
Observe that $(A+B)(A-B) = A(A-B) + B(A-B) = A^2 - AB + BA - B^2$. Now choose $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$. Notice that in this case we have $A^2 = BA = A$ and $B^2 = AB = B$. And so $(A+B)(A-B) = A - B + A - B = 2A - 2B$ while $A^2 - B^2 = A - B$. But it is clear that for our choice of A and B it is not true that $2A - 2B = A - B$ and so we have a counter example.

d. If $AB = CB$ and B is nonsingular, then $A = C$.

I'm assuming that I don't know that if B is nonsingular then it has a multiplicative inverse. Going with what we know from the book, if B is nonsingular then $B\mathbf{x} = \mathbf{b}$ has a solution and that solution is unique.

I wish to construct a series of vectors from the columns of I_n. We choose some $\mathbf{b}_i$ such that $b_i = 1$ for some $1 \leq i \leq n$ and $b_j = 0$ for every $j \neq i$ and $1 \leq j \leq n$. We know we can find some unique $\mathbf{x}_i$ such that $B\mathbf{x}_i = \mathbf{b}_i$.

Now we form a unique matrix $B^{-1} = \begin{bmatrix} \mathbf{x}_1 & \dots & \mathbf{x}_n \end{bmatrix}$. Notice that $BB^{-1} = \begin{bmatrix} B\mathbf{x}_1 & \dots & B\mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{b_1} & \dots & \mathbf{b}_n \end{bmatrix} = I_n$. So now that we have established that for every B there exists some B^{-1} such that $BB^{-1} = I_n$ this problem becomes trivial. In fact $AB = CB$ implies that $ABB^{-1} = CBB^{-1}$ and so obviously $A = C$. $\square$

e. If $AB = BC$ and B is nonsingular then $A = C$

We choose $B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$. If $B\mathbf{x} = \mathbf{0}$ then $x_1 - x_2 = 0$ and $x_1 + 0 \cdot x_2 = 0$. It quickly follows that $\mathbf{x} = (0, 0)$ and so B is nonsingular. Now if $A = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ then $AB = BC = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ but $A \neq C$

2.1.7 Find all 2×2 matrices $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ satisfying

$$A^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 = \begin{bmatrix} a^2 + bc & ab + bd \\ ca + dc & cb + d^2 \end{bmatrix} = \begin{bmatrix} a^2 + bc & b(a + d) \\ c(a + d) & cb + d^2 \end{bmatrix}$$

And so for each of the following cases we know that either $a = -d$ or $b = c = 0$

a. $A^2 = I_2$

If $b = c = 0$ then $a^2 = d^2 = 1$. On the other hand if $a = -d$ then $a^2 + bc = d^2 + bc = 1$ or $1 - a^2 = 1 - d^2 = bc$. If $a^2 = 1$ then either $b = 0$ or $c = 0$. Otherwise $b \neq 0$ and $c = \frac{1-a^2}{b}$. So then our cases are

$$\begin{bmatrix} a & 0 \\ 0 & \pm a \end{bmatrix}, \begin{bmatrix} a & b \\ 0 & -a \end{bmatrix}, \begin{bmatrix} a & 0 \\ c & -a \end{bmatrix} \text{ with } a \in \{1, -1\} \text{ and } b, c \in \mathbb{R}$$

$$\text{or} \begin{bmatrix} a & b \\ \frac{1-a^2}{b} & -a \end{bmatrix} \text{ with } a, b \in \mathbb{R} \text{ and } b \neq 0$$

b. $A^2 = O$

We know that $a^2 + bc = d^2 + bc = 0$ and so if $b = c = 0$ then $a = d = 0$ implying that $a = -d$, so we don't need a special case for this. Now assume that $a = -d$ then $a^2 = -bc$. Now if $b = 0$ then $a^2 = d^2 = 0$. If $b \neq 0$ then $c = -\frac{a^2}{b}$. So our cases are

$$\begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix} \text{ with } c \in \mathbb{R}$$

$$\text{or} \begin{bmatrix} a & b \\ -\frac{a^2}{b} & -a \end{bmatrix} \text{ with } a, b \in \mathbb{R} \text{ and } b \neq 0$$

c. $A^2 = -I_2$

We wish for $a^2 + bc = -1$ and so if either $b = 0$ or $c = 0$ then $a^2 = -1$. We are restricting ourselves to real numbers for this exercise and so we will say that $b, c \in \mathbb{R} \setminus \{0\}$ and $a = -d$. Then $a^2 + bc = -1$ or $c = -\frac{1+a^2}{b}$. Thus A will take the form

$$\begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix} \text{ with } c \in \mathbb{R}$$

$$\text{or} \begin{bmatrix} a & b \\ -\frac{a^2}{b} & -a \end{bmatrix} \text{ with } a, b \in \mathbb{R} \text{ and } b \neq 0$$

2.1.8 For each of the following matrices A, find a formula for A^k for positive integers k.

a. $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

We shall prove this using induction. First we must form our basis by observing that $A^2 = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}$. Now lets assume that $A^{k-1} = \begin{bmatrix} 2^{k-1} & 0 \\ 0 & 3^{k-1} \end{bmatrix}$ for any $k > 2$. We see that $A^k = A^{k-1}A = \begin{bmatrix} 2^{k-1} & 0 \\ 0 & 3^{k-1} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix}$.

b. $A = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$

I'm assuming that this notation means that $a_{ij} = 0$ if $i \neq j$ and $a_{ij} = d_i$ if $i = j$. We proceed using induction. First we must form our basis by observing that $\text{ent}_{ij}(A^2) = \sum_{p=1}^{n} a_{ip}a_{pj}$. Now if $i \neq j$ then either $i \neq p$ or $j \neq p$. Thus for all p we will have $a_{ip} = 0$ or $a_{pj} = 0$ and so $\text{ent}_{ij}(A^2) = 0$. Now if $i = j$ we have $a_{ip}a_{pj} = 0$ for all $p \neq i$ and $a_{ip}a_{pj} = d_i^2$ for all $p = i$. Thus $\text{ent}_{ij}(A^k) = d_i^2$ for all $i = j$. This means that

$$A^2 = \begin{bmatrix} {d_1}^2 & & & \\ & {d_2}^2 & & \\ & & \ddots & \\ & & & {d_n}^2 \end{bmatrix}$$

Now let us assume that

$$A^{k-1} = \begin{bmatrix} {d_1}^{k-1} & & & \\ & {d_2}^{k-1} & & \\ & & \ddots & \\ & & & {d_n}^{k-1} \end{bmatrix} \text{ for } k > 2$$

We see that $A^k = A^{k-1}A$ and $\text{ent}_{ij}(A^k) = \sum_{p=1}^{n} \text{ent}_{ip}(A^{k-1})a_{pj}$. Now if $i \neq j$ then either $i \neq p$ or $j \neq p$. Thus for all p we will have $\text{ent}_{ip}(A^{k-1})a_{pj} = 0$. Now if $i = j$ we have $\text{ent}_{ip}(A^{k-1})a_{pj} = 0$ for all $p \neq i$ and $\text{ent}_{ip}(A^{k-1})a_{pj} = d_i^{k-1}d_i = d_i^k$ for all $p = i$. Thus $\text{ent}_{ij}(A^k) = d_i^k$ for all $i = j$. This means that

$$A^k = \begin{bmatrix} {d_1}^k & & & \\ & {d_k}^2 & & \\ & & \ddots & \\ & & & {d_n}^k \end{bmatrix}$$

c. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

It is easy to see that $A^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Now if $k > 2$ and we assume that $A^{k-1} = \begin{bmatrix} 1 & k-1 \\ 0 & 1 \end{bmatrix}$ then we see that $A^k = A^{k-1}A = \begin{bmatrix} 1 & k-1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$. And so induction tells us that this must be the case for all $k > 0$

2.1.11 a. Suppose $A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times m}$ and $BA = I_n$. Prove that if for some $\mathbf{b} \in \mathbb{R}^m$ the equation $A\mathbf{x} = \mathbf{b}$ has a solution, then that solution is unique.

Multiplying both sides of the equation by B gives us $B(A\mathbf{x}) = B\mathbf{b}$. We can then apply theorem 9 to obtain $B(A\mathbf{x}) = (BA)\mathbf{x} = I_n\mathbf{x} = \mathbf{x} = B\mathbf{b}$. Because matrix multiplication is well defined, we can assume that $\mathbf{x}$ is unique.

b. Suppose $A \in \mathcal{M}_{m\times n}, C \in \mathcal{M}_{n\times m}$ and $AC = I_m$. Prove that the system $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b} \in \mathbb{R}^m$.

If we let $\mathbf{x} = C\mathbf{b}$ then with theorem 9 we have $A(C\mathbf{b}) = (AC)\mathbf{b} = I_m\mathbf{b} = \mathbf{b}$. We see that we can always find at least one solution, so the system is consistent.

c. Suppose $A \in \mathcal{M}_{m\times n}$ and $B, C \in \mathcal{M}_{n\times m}$ are matrices that satisfy $BA = I_n$ and $AC = I_m$. Prove that $B = C$.

This is mostly just using the properties of identities and associative properties of multiplication. Witness

$$B = BI_m = B(AC) = (BA)C = I_nC = C$$

2.1.12 An $n \times n$ matrix is called a *permutation matrix* if it has a single 1 in each row and column and all its remaining entries are 0.

a. Write down all the 2×2 permutation matrices. How many are there?

The easiest thing is to answer how many $n \times n$ permutation matrices there are. In the first row there are n choices for where we place our 1. In the second row we have n choices minus the column we put our 1 for the first row. In this way we see that we have $n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1 = n!$ ways of making a permutation matrix. So we have 2 possible 2×2 matrices.

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

b. Write down all the 3×3 permutation matrices. How many are there?

We already know there are $3! = 6$ of these.

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

c. Show that the product of two permutation matrices is again a permutation matrix. Do they commute?

We take two $n \times n$ permutation matrices and label them P_1 and P_2 without loss of generality. If we multiply them then we get $P_1P_2 = \begin{bmatrix} P_1\mathbf{c}_1(P_2) & \dots & P_1\mathbf{c}_n(P_2) \end{bmatrix}$. If $p_{i_j j}$ is the entry in column j of P_2 which is equal to 1, then $P_1P_2 = \begin{bmatrix} \mathbf{c}_{i_1}(P_1) & \dots & \mathbf{c}_{i_n}(P_1) \end{bmatrix}$. Now because P_2 contains only one 1 in each row and column, we know that if $j \neq k$ then $i_j \neq i_k$. Thus P_2 simply shuffles the columns of P_1. Since rearranging the columns of a matrix with one 1 in each row and column will still leave us with a matrix that has one 1 in each row and column, then P_1P_2 is a permutation matrix. $\square$ The permutations may commute but they do not necessarily. For example, if one of the matrices is I_n then they will commute. However we can easily come up with a counterexample.

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

e. If A is an $n \times n$ matrix and P is an $n \times n$ permutation matrix, describe the columns of AP and the rows of PA.

We choose some column $\mathbf{c}_j(P)$. Let i be the index such that $p_{ij} = 1$ with and all other entries in $\mathbf{c}_j = 0$. Now $\mathbf{c}_j(AP) = \mathbf{c}_i(A)$. Thus P permutes the columns of A when multiplied on the right. Similarly we choose some row $\mathbf{r}_i(P)$ and let j be the index such that $p_{ij} = 1$. Now $\mathbf{r}_i(PA) = \mathbf{r}_j(A)$. Thus P permutes the rows of A when multiplied on the left.

2.1.14 Find all 2×2 matrices A that commute with all 2×2 matrices B. That is, if $AB = BA$ for all $B \in \mathcal{M}_{2\times 2}$, what are the possible matrices that A can be?

$$AB = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = BA$$
$$\begin{bmatrix} a_1b_1 + a_2b_3 & a_1b_2 + a_2b_4 \\ a_3b_1 + a_4b_3 & a_3b_2 + a_4b_4 \end{bmatrix} = \begin{bmatrix} b_1a_1 + b_2a_3 & b_1a_2 + b_2a_4 \\ b_3a_1 + b_4a_3 & b_3a_2 + b_4a_4 \end{bmatrix}$$

Looking at the top left entries of both sides of the equation leads us to $a_2b_3 = b_2a_3$. However the only solution to this equation for a_2 and a_3 that does not depend on b_2 or b_3 is $a_2 = a_3 = 0$. This vastly simplifies the equation we are dealing with.

$$\begin{bmatrix} a_1b_1 & a_1b_2 \\ a_4b_3 & a_4b_4 \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_4b_2 \\ a_1b_3 & a_4b_4 \end{bmatrix}$$

So we need to find solutions to $a_1b_2 = a_4b_2$ and $a_4b_3 = a_1b_3$. Both of these equations are always satisfied only if $a_1 = a_4$. And so we have our solution. $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \quad \forall a \in \mathbb{R}$