More on matrix operations

Special cases of matrix multiplication:

row vector · (Column) Vector:

Let $\underline{a} = \begin{bmatrix} 1 & -2 & 7 \end{bmatrix}$ $\underline{b} = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}$. Then

$\underline{a}\,\underline{b}$: $(1\times \underbrace{3)(3}_{\text{match ✓}}\times 1) \longrightarrow$ 1×1 result.

(1, 1)-entry is $1(0) + (-2)(3) + 7(2) = 8$,
so result is:

$$\begin{bmatrix} 1 & -2 & 7 \end{bmatrix}\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \end{bmatrix}$$

---

matrix · vector: Let

$$\underline{A} = \begin{bmatrix} 1 & 7 & 2 \\ 0 & -1 & 0 \\ 3 & 4 & 5 \end{bmatrix} \qquad \underline{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$\underline{A}\,\underline{x}$: $(3\times \underbrace{3)(3}_{\text{match ✓}}\times 1) \longrightarrow$ result is 3×1

The result is

$$\underline{A}\,\underline{x} = \begin{bmatrix} 1\cdot 1 + 7\cdot 0 + 2\cdot(-1) \\ 0\cdot 1 + (-1)\cdot 0 + 0\cdot(-1) \\ 3\cdot 1 + 4\cdot 0 + 5\,(-1) \end{bmatrix}_{(3\times 1)}$$

$$= \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix}$$

Notice that matrix · vector resulted in a vector of the same size. This always happens.

Writing systems of linear equations as matrix equations (and vice-versa)

Let $\underline{A} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 7 \\ 2 & 7 & 9 \end{bmatrix}$, $\underline{b} = \begin{bmatrix} 4 \\ 20 \\ 23 \end{bmatrix}$, and

$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, where x_1, x_2, x_3 are unknowns.

(so $\underline{x}$ is an unknown vector). Then the equality

$$\underline{A}\,\underline{x} = \underline{b}, \text{ i.e. } \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 7 \\ 2 & 7 & 9 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 20 \\ 23 \end{bmatrix},$$

is valid if and only if

$$\begin{bmatrix} x_1 + 2x_2 + x_3 \\ 3x_1 + 8x_2 + 7x_3 \\ 2x_1 + 7x_2 + 9x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 20 \\ 23 \end{bmatrix},$$

which is valid if and only if all 3 entries are equal,

i.e. $$\begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 3x_1 + 8x_2 + 7x_3 = 20 \\ 2x_1 + 7x_2 + 9x_3 = 23 \end{cases}.$$

That is, given $\underline{A}$ and $\underline{b}$ above, the vector $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is a solution to the equation

$$\underline{A}\,\underline{x} = \underline{b} \quad \left(\text{a.k.a. } \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 7 \\ 2 & 7 & 9 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 20 \\ 23 \end{bmatrix}\right)$$

if and only if x_1, x_2, x_3 are a solution to the system of linear equations

$$\begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 3x_1 + 8x_2 + 7x_3 = 20 \\ 2x_1 + 7x_2 + 9x_3 = 23 \end{cases}$$

In summary: (solutions to systems of linear equations) correspond to ((vector) solutions to matrix equations)

Rules of matrix operations:

The following equations / "properties" are valid for matrices (notice they are mostly the same as the properties of basic algebra for numbers). Throughout, assume $\underline{A}$, $\underline{B}$, and $\underline{C}$ are matrices with matching dimensions (say $n \times n$)

1) (Commutative addition): $\underline{A} + \underline{B} = \underline{B} + \underline{A}$
2) (Associative addition): $(\underline{A}+\underline{B})+\underline{C} = \underline{A}+(\underline{B}+\underline{C})$
3) (associative multiplication): $\underline{A}(\underline{B}\,\underline{C}) = (\underline{A}\,\underline{B})\underline{C}$
4) (Distributive laws): $\underline{A}(\underline{B}+\underline{C}) = \underline{A}\,\underline{B} + \underline{A}\,\underline{C}$
   and $(\underline{A}+\underline{B})\underline{C} = \underline{A}\,\underline{C} + \underline{B}\,\underline{C}$.

— — — — — —

(Stuff involving scalars) Let α and β be scalars.

Scalars always "commute" with matrices and vectors, e.g. $3\underline{A} = \underline{A} \cdot 3$. In class we had an example of $\alpha\begin{bmatrix}2\\3\end{bmatrix} + \beta\begin{bmatrix}0\\\pi\end{bmatrix} = \begin{bmatrix}2\alpha\\3\alpha+\beta\pi\end{bmatrix}$.

We generally write (scalar)(vector), but it is totally valid to write

$$\begin{bmatrix}2\\3\end{bmatrix}\alpha + \begin{bmatrix}0\\\pi\end{bmatrix}\beta \quad \left(\text{which is again } = \begin{bmatrix}2\alpha\\3\alpha+\beta\pi\end{bmatrix}\right).$$

This is useful when multiply (matrices)(vectors):

Example:

Let $\underline{A} = \begin{bmatrix} 2 & 7 \\ -1 & 0 \end{bmatrix}$ and $\underline{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Then $\underline{A}\,\underline{x} = \begin{bmatrix} 2\cdot 1 + 7\cdot(-1) \\ (-1)\cdot 1 + 0\cdot(-1) \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$

If we now wanted to compute

$\underline{A}(3\cdot\underline{x})$, we get $\begin{bmatrix} 2 & 7 \\ -1 & 0 \end{bmatrix}\left(\begin{bmatrix} 3 \\ -3 \end{bmatrix}\right)$

$$= \begin{bmatrix} 2 & 7 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

$$= \begin{bmatrix} 2\cdot 3 + 7\cdot(-3) \\ (-1)\cdot 3 + 0\cdot(-3) \end{bmatrix}$$

$$= \begin{bmatrix} -15 \\ -3 \end{bmatrix}$$

$$= 3\begin{bmatrix} -5 \\ -1 \end{bmatrix}$$

$$= 3(\underline{A}\,\underline{x}),$$

i.e. $\underline{A}(3\underline{x}) = 3(\underline{A}\,\underline{x})$.

Warning: When multiplying numbers a and b, we know $ab = ba$. When multiplying matrices, we cannot assume that $\underline{A}\,\underline{B} = \underline{B}\,\underline{A}$.

Ex: $\underline{A} = \begin{bmatrix} 2 & -1 \\ -4 & 3 \end{bmatrix}$ $\underline{B} = \begin{bmatrix} 1 & 5 \\ 3 & 7 \end{bmatrix}$.

Then $\underline{A}\,\underline{B} = \begin{bmatrix} -1 & 3 \\ 5 & 1 \end{bmatrix}$, but

$\underline{B}\,\underline{A} = \begin{bmatrix} -18 & 14 \\ -22 & 18 \end{bmatrix}$,

so $\underline{A}\,\underline{B} \neq \underline{B}\,\underline{A}$.

Ex in class we had

$$\underline{A} = \begin{bmatrix} 1 & 0 \\ 3 & -1 \\ 2 & 4 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix},$$

and found that

$$\underline{A}\,\underline{B} = \begin{bmatrix} -1 & 2 \\ -4 & 5 \\ 2 & 8 \end{bmatrix}.$$

However, $\underline{B}\,\underline{A} : (2\times\underbrace{2)(3}_{\text{mismatch!}}\times 2)$ isn't even defined, much less equal to $\underline{A}\,\underline{B}$.

A few other bits of matrix notation and facts

$\mathbf{AB} \neq \mathbf{BA}$. Other exceptions are associated with zero matrices. A **zero matrix** is one whose elements are *all* zero, such as

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We ordinarily denote a zero matrix (whatever its size) by $\mathbf{0}$. It should be clear that for any matrix $\mathbf{A}$,

$$\mathbf{0} + \mathbf{A} = \mathbf{A} = \mathbf{A} + \mathbf{0}, \quad \mathbf{A0} = \mathbf{0}, \quad \text{and} \quad \mathbf{0A} = \mathbf{0},$$

where in each case $\mathbf{0}$ is a zero matrix of appropriate size. Thus zero matrices appear to play a role in the arithmetic of matrices similar to the role of the real number 0 in ordinary arithmetic.

Recall that an *identity matrix* is a square matrix $\mathbf{I}$ that has ones on its principal diagonal and zeros elsewhere. Identity matrices play a role in matrix arithmetic which is strongly analogous to that of the real number 1, for which $a \cdot 1 = 1 \cdot a = a$ for all values of the real number a. For instance, you can check that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Similarly, if

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \text{and} \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

then $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$. For instance, the element in the second row and third column of $\mathbf{AI}$ is

$$(a_{21})(0) + (a_{22})(0) + (a_{23})(1) = a_{23}.$$

Recall that the $n \times n$ **identity matrix** is the diagonal matrix

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{1}$$

having ones on its main diagonal and zeros elsewhere. It is not difficult to deduce directly from the definition of the matrix product that $\mathbf{I}$ acts like an identity for matrix multiplication:

$$\mathbf{AI} = \mathbf{A} \quad \text{and} \quad \mathbf{IB} = \mathbf{B} \tag{2}$$

if the sizes of $\mathbf{A}$ and $\mathbf{B}$ are such that the products $\mathbf{AI}$ and $\mathbf{IB}$ are defined. It is, nevertheless, instructive to derive the identities in (2) formally from the two basic facts about matrix multiplication that we state below. First, recall that the notation

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \cdots & \mathbf{a}_n \end{bmatrix} \tag{3}$$

expresses the $m \times n$ matrix $\mathbf{A}$ in terms of its column vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots, \mathbf{a}_n$.

Fact 1 $\quad$ $\mathbf{Ax}$ *in terms of columns of* $\mathbf{A}$

If $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is an n-vector, then

$$\mathbf{Ax} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n. \tag{4}$$

The reason is that when each row vector of $\mathbf{A}$ is multiplied by the column vector $\mathbf{x}$, its jth element is multiplied by x_j.

Fact 2 $\quad$ $\mathbf{AB}$ *in terms of columns of* $\mathbf{B}$

If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix}$ is an $n \times p$ matrix, then

$$\mathbf{AB} = \begin{bmatrix} \mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_p \end{bmatrix}. \tag{5}$$

That is, *the* j*th column of* $\mathbf{AB}$ *is the product of* $\mathbf{A}$ *and the* j*th column of* $\mathbf{B}$. The reason is that the elements of the jth column of $\mathbf{AB}$ are obtained by multiplying the individual rows of $\mathbf{A}$ by the jth column of $\mathbf{B}$.

Example 1 The third column of the product $\mathbf{AB}$ of the matrices

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ 4 & 0 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 3 & 7 & 5 & -4 \\ -2 & 6 & 3 & 6 \\ 5 & 1 & -2 & -1 \end{bmatrix}$$

is

$$\mathbf{Ab}_3 = \begin{bmatrix} 2 & -1 & 0 \\ 4 & 0 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 7 \\ 14 \end{bmatrix}.$$

■

To prove that $\mathbf{AI} = \mathbf{A}$, note first that

$$\mathbf{I} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix}, \tag{6}$$

where the jth column vector of $\mathbf{I}$ is the jth **basic unit vector**

$$\mathbf{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j\text{th entry}. \tag{7}$$

If $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$, then Fact 1 yields

$$\mathbf{Ae}_j = 0 \cdot \mathbf{a}_1 + \cdots + 1 \cdot \mathbf{a}_j + \cdots + 0 \cdot \mathbf{a}_n = \mathbf{a}_j. \tag{8}$$

Hence Fact 2 gives

$$\begin{aligned} \mathbf{AI} &= \mathbf{A}\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{Ae}_1 & \mathbf{Ae}_2 & \cdots & \mathbf{Ae}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}; \end{aligned}$$

that is, $\mathbf{AI} = \mathbf{A}$. The proof that $\mathbf{IB} = \mathbf{B}$ is similar. (See Problems 41 and 42.)