# Section 3: Linear Algebra Review

The branch of mathematics called *linear algebra* mainly concerns solving a *linear system of equations*. Linear algebra is used in a variety of problems in robotics, some of which we will discuss in later sections. Here, we review some basic concepts.

By *linear* we mean that the variables (the unknowns in the equation) are only multiplied by constants. For example, if we have three variables $x_1$, $x_2$ and $x_3$ and three constants $a_1$, $a_2$ and $a_3$ then the following terms are linear:

- $x_1$
- $x_2 + x_1 + x_2$
- $a_1^2 + x_2/a_2$
- $(1 - x_1)a_3 a_2$

On the other hand, the following *nonlinear* terms involve two variables multiplied by each other, trigonometric, exponential, or other functions of the variables that cannot be expressed in linear form:

- $x_1^2$
- $x_1(1 + x_2)$
- $\cos(x_2)$
- $\sqrt{x_3}$

A linear system of $m$ equations with $n$ variables $\{x_1, x_2, \cdots, x_n\}$ is written in *standard form* as

$$\begin{aligned}
\text{Equation 1}: &\quad a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
\text{Equation 2}: &\quad a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
&\quad \vdots \\
\text{Equation } m: &\quad a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m
\end{aligned}$$

where $a_{ij}$ and $b_i$ are coefficients in which the index $i$ indicates the equation number (from 1 to $m$) and the index $j$ indicates the corresponding variable $x_j$ (from 1 to $n$). In standard form, the right hand side (RHS) contains all of the constants (that do not multiply any variables) and the variables, and their coefficients, on the left hand side (LHS) are summed in increasing order from $x_1$ to $x_n$.

A linear system of equations in standard form can be written in *matrix form* as $\mathbf{Ax} = \mathbf{b}$. In this equation, the boldface $\mathbf{A}$ is the matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

that contains the coefficients of the Equations 1 through $m$ given above. A matrix is an array of numbers arranged in a table form (i.e., as a grid, or spreadsheet) with rows and columns. We say that a matrix is of size $(m \times n)$ (pronounced "m by n") to indicate that there are $m$ rows and $n$ columns. An element of the matrix $\mathbf{A}$ is referred to using index notation as $a_{ij}$ to indicate that it is the element corresponding to the $i$-th row and the $j$-th column. The element $a_{ij}$ corresponds to the coefficient that multiplies the $x_j$ term in the $i$-th equation. In the equation $\mathbf{Ax} = \mathbf{b}$ the symbol $\mathbf{x}$ represents a column vector (a matrix with one column and $n$ rows) that lists the $n$ variables of the linear system

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and the symbol $\mathbf{b}$ represents a column vector of the $m$ constants on the RHS of the linear system in standard form

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

We will generally assume that, unless otherwise specified, vectors in this course are column vectors. Together, the equation $\mathbf{Ax} = \mathbf{b}$ is written

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\mathbf{b}} \tag{1}$$

---

**Example: Converting a linear system of equation into matrix form**
Suppose we have the linear system of equations

$$\begin{aligned} 1 + 3x_2 &= 0 \\ 6x_2 + x_1 + 2x_3 &= 2 \\ -5x_3 + 4x_1 &= 1 \end{aligned}$$

that we wish to convert into matrix form. First, we move the constants to the RHS and rearrange the variables in increasing order (i.e., so that the $x_2$ term follows the $x_1$ term, and so on)

$$\begin{aligned} 0x_1 + 3x_2 + 0x_3 &= -1 \\ 1x_1 + 6x_2 + 2x_3 &= 2 \\ 4x_1 + 0x_2 - 5x_3 &= 1 \end{aligned}$$

All variables have been written with coefficients (including ones and zeros) for clarity. The coefficient of the LHS are then written in matrix form:

$$\mathbf{A} = \begin{pmatrix} 0 & 3 & 0 \\ 1 & 6 & 2 \\ 4 & 0 & -5 \end{pmatrix}$$

the column vector of variables is

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

and the constant vector on the RHS is

$$\mathbf{b} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

Putting everything together, the system of equations in matrix form $\mathbf{Ax} = \mathbf{b}$ is

$$\underbrace{\begin{pmatrix} 0 & 3 & 0 \\ 1 & 6 & 2 \\ 4 & 0 & -5 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}}_{\mathbf{b}}$$

---

To convert an equation in matrix form (such as Eq. 1) into standard form we reconstruct the $j$-th equation by multiplying the $j$-th row of the matrix $\mathbf{A}$ with $\mathbf{x}$ and equation to the $j$-th constant in $\mathbf{b}$. This "multiplication" of a row-vector by a column-vector is similar to the familiar dot product. Thus, the $j$-th equation is obtained as follows

$$\begin{pmatrix} a_{j1} & a_{j2} & \cdots & a_{jn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_3 \end{pmatrix} = a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j$$

Recall that there are $m$ equations, so repeating this for $j = \{1, 2, \cdots, m\}$ we can obtain the original system of equations in standard form.

---

**Example: Converting an equation in matrix form into standard form**

Suppose we begin with the system

$$\underbrace{\begin{pmatrix} 0 & 3 & 0 \\ 1 & 6 & 2 \\ 4 & 0 & -5 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}}_{\mathbf{b}}$$

The first equation is obtained by multiplying the first row of $\mathbf{A}$ by $\mathbf{x}$ and equating this to the first element of $\mathbf{b}$

$$\begin{pmatrix} 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0x_1 + 3x_2 + 0x_3 = -1$$

Similarly, for the second equation

$$\begin{pmatrix} 1 & 6 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 1x_1 + 6x_2 + 2x_3 = 2$$

and for the third

$$\begin{pmatrix} 4 & 0 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4x_1 + 0x_2 - 5x_3 = 1$$

---

The transpose of a matrix is denoted with the superscript symbol $^{\mathrm{T}}$ and is obtained by interchanging the rows and columns of the matrix. Thus, the transpose of the matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

is

$$\mathbf{A}^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

---

**Example: Transposing matrices and vectors**

The "skinny" matrix $\mathbf{A}$ becomes a "wide" matrix $\mathbf{A}^{\mathrm{T}}$ after transposing:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 7 & -2 \\ -1 & 3 \\ 9 & 5 \end{pmatrix} \qquad \mathbf{A}^{\mathrm{T}} = \begin{pmatrix} 0 & 7 & -1 & 9 \\ 1 & -2 & 3 & 5 \end{pmatrix}$$

Similarly, the column vector becomes a row vector after transposing

$$\mathbf{b} = \begin{pmatrix} 4 \\ 3 \\ -1 \\ 10 \end{pmatrix} \qquad \mathbf{b}^{\mathrm{T}} = \begin{pmatrix} 4 & 3 & -1 & 10 \end{pmatrix}$$

For a square matrix (with an equal number of rows and columns) the diagonal elements remain unchanged when transposing. In the following example, the elements on the diagonal $(0, 2, 3, 1)$ remain unchanged:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 5 \\ 7 & 2 & 0 & 3 \\ -1 & 3 & 3 & -3 \\ 9 & 5 & -1 & 1 \end{pmatrix} \qquad \mathbf{A}^{\mathrm{T}} = \begin{pmatrix} 0 & 7 & -1 & 9 \\ 1 & 2 & 3 & 5 \\ 1 & 0 & 3 & -1 \\ 5 & 3 & -3 & 1 \end{pmatrix}$$

---

Matrix addition and subtraction requires that both matrices are of the same size and the operation happens element-wise. For example, $\mathbf{C} = \mathbf{A} + \mathbf{B}$ implies that $\mathbf{A}$ and $\mathbf{B}$ are the same size and each element of $\mathbf{C}$ is obtained by adding the corresponding element of $\mathbf{A}$ and $\mathbf{B}$, that is $c_{ij} = a_{ij} + b_{ij}$. In general, if

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

then

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix} = \begin{pmatrix} (a_{11}+b_{11}) & (a_{12}+b_{12}) & \cdots & (a_{1n}+b_{1n}) \\ (a_{21}+b_{21}) & (a_{22}+b_{22}) & \cdots & (a_{2n}+b_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ (a_{m1}+b_{m1}) & (a_{m2}+b_{m2}) & \cdots & (a_{mn}+b_{mn}) \end{pmatrix}$$

And similarly for subtraction:

$$\mathbf{C} = \mathbf{A} - \mathbf{B}$$

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} - \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix} = \begin{pmatrix} (a_{11}-b_{11}) & (a_{12}-b_{12}) & \cdots & (a_{1n}-b_{1n}) \\ (a_{21}-b_{21}) & (a_{22}-b_{22}) & \cdots & (a_{2n}-b_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ (a_{m1}-b_{m1}) & (a_{m2}-b_{m2}) & \cdots & (a_{mn}-b_{mn}) \end{pmatrix}$$

Scalar multiplication also occurs element-wise. If the matrix $\mathbf{A}$ is multiplied by a scalar $\alpha$ then the result is obtained by multiplying each element $a_{ij}$ by $\alpha$. Thus,

$$\alpha\mathbf{A} = \alpha \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{pmatrix}$$

---

**Example: Matrix addition, subtraction and scalar multiplication**
Let

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 7 & 2 & 0 \\ -1 & 3 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 3 & 4 & 6 \\ -1 & 11 & 8 \\ 2 & 5 & 1 \end{pmatrix}$$

then by adding these two matrices we obtain

$$\mathbf{A}+\mathbf{B}=\begin{pmatrix}(0+3) & (1+4) & (1+6)\\(7-1) & (2+11) & (0+8)\\(-1+2) & (3+5) & (3+1)\end{pmatrix}=\begin{pmatrix}3 & 5 & 7\\6 & 13 & 8\\1 & 8 & 4\end{pmatrix}$$

Similarly, subtracting these matrices

$$\mathbf{A}-\mathbf{B}=\begin{pmatrix}(0-3) & (1-4) & (1-6)\\(7+1) & (2-11) & (0-8)\\(-1-2) & (3-5) & (3-1)\end{pmatrix}=\begin{pmatrix}-3 & -3 & -5\\8 & -9 & -8\\-3 & -2 & 2\end{pmatrix}$$

Last, multiplying $\mathbf{A}$ by a scalar $\alpha = 2$ gives

$$2\mathbf{A}=2\begin{pmatrix}0 & 1 & 1\\7 & 2 & 0\\-1 & 3 & 3\end{pmatrix}=\begin{pmatrix}0 & 2 & 2\\14 & 4 & 0\\-2 & 6 & 6\end{pmatrix}$$

---

To multiply a matrix $\mathbf{A}$ by a matrix $\mathbf{B}$ their respective sizes must be compatible. Specifically, for the operation $\mathbf{C} = \mathbf{AB}$ to be well defined, the number of columns of $\mathbf{A}$ should be equal to the number of rows of $\mathbf{B}$. This condition is satisfied if $\mathbf{A}$ is size $(m \times n)$ and $\mathbf{B}$ is size $(n \times p)$. The result of the multiplication will be a new matrix $\mathbf{C}$ of size $(m \times p)$.

A convenient way to ensure a matrix operation is valid, and to determine the size of the resulting matrix, is to write out the respective sizes of each matrix, side-by-side, and cross out the "cancelling terms". For example, in the above case we can write

$$(m\cancel{\times n)(n} \times p) \rightarrow (m \times p)$$

and strike out the two $n$ terms since they "cancel". This is particularly useful for a complex expression involving several matrices. Suppose that in addition to $\mathbf{A}$ and $\mathbf{B}$ as defined above, we have a $(n \times q)$ sized matrix $\mathbf{Q}$, a $(q \times r)$ sizes matrix $\mathbf{R}$ and a $(r \times n)$ sized matrix $\mathbf{M}$. The equation $\mathbf{D} = \mathbf{QRM}$ is well defined because

$$(n\cancel{\times q)(q \times r)(r} \times n) \rightarrow (n \times n)$$

which implies the result $\mathbf{D}$ is a square matrix with $n$ rows and $n$ columns. The operation $\mathbf{C} = \mathbf{ADB}$ is also valid since

$$(m\cancel{\times n)(n \times n)(n} \times p) \rightarrow (m \times p)$$

While the multiplication $\mathbf{AD}$ is valid since $\mathbf{A}$ has the same number of columns as $\mathbf{B}$ has rows, the multiplication in reverse order $\mathbf{DA}$ is not valid. In general, matrix multiplication is not *commutative* (i.e., the order of matrix multiplication matters). For clarity, we will someitmes say $\mathbf{A}$ *pre-multiplies* $\mathbf{B}$ to denote that we are referring to $\mathbf{AB}$ instead of $\mathbf{BA}$.

---

**Example: Determining the size of a matrix resulting from a matrix multiplication**

Let

$$\mathbf{A}=\begin{pmatrix}0 & 1 & 1\\7 & 2 & 0\end{pmatrix}, \quad \mathbf{B}=\begin{pmatrix}9 & 1 & 3\\1 & 2 & 1\\-5 & 0 & 10\end{pmatrix}, \quad \text{and} \quad \mathbf{C}=\begin{pmatrix}3 & 2\\-1 & 8\\0 & 5\end{pmatrix},$$

The size of **A** is $(2 \times 3)$, the size of **B** is $(3 \times 3)$, the size of **C** is $(3 \times 2)$. Thus, the matrix $D = \mathbf{ABC}$ is of size

$$(2 \cancel{\times 3)(3 \times 3)(3} \times 2) \rightarrow (2 \times 2)$$

Note that the following operations are valid **AB**, **AC**, **CA**, **BC** whereas the following are not **BA**, **CB**.

---

After determining the size of the matrix resulting from a matrix multiplication we can sketch out the result using index notation. For example, after multiplying **A** of size $(m \times n)$ by **B** of size $(n \times p)$ we can sketch out the result $\mathbf{C} = \mathbf{AB}$ of size $(m \times p)$ as follows:

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix}$$

Each element $c_{ij}$ is a result of multiplying the $i$-th row of **A** by the $j$-th column of **B**

$$\begin{aligned} c_{ij} &= \begin{pmatrix} i\text{-th row of } \mathbf{A} \end{pmatrix} \begin{pmatrix} j\text{-th} \\ \text{column} \\ \text{of} \\ \mathbf{B} \end{pmatrix} \\ &= \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} \\ &= (a_{i1}b_{1j}) + (a_{i2}b_{2j}) + \cdots + (a_{in}b_{nj}) \end{aligned}$$

Then by evaluating for all of the values $c_{ij}$ we arrive at the resulting matrix **C**.

---

**Example: Evaluating the matrix resulting from a matrix multiplication**
Returning to our previous example, let

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 7 & 2 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 9 & 1 & 3 \\ 1 & 2 & 1 \\ -5 & 0 & 10 \end{pmatrix}, \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} 3 & 2 \\ -1 & 8 \\ 0 & 5 \end{pmatrix},$$

The product $\mathbf{D} = \mathbf{AB}$ is of size $(2 \times 3)$ and writen in index notation as

$$\mathbf{D} = \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \end{pmatrix}$$

we can evaluate each element as follows:

$$d_{11} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 9 \\ 1 \\ -5 \end{pmatrix} = 0(9) + 1(1) + 1(-5) = -4$$

$$d_{12} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} \begin{pmatrix} b_{12} \\ b_{22} \\ b_{32} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 0(1) + 1(2) + 1(0) = 2$$

Repeating this procedure for the remaining indices ($d_{13}$, $d_{21}$, $d_{22}$, and $d_{23}$) we arrive at the final result

$$\mathbf{D} = \begin{pmatrix} -4 & 2 & 11 \\ 65 & 11 & 23 \end{pmatrix}$$

If a matrix multiplicaton involves several matrices, we can iterate this process for every pair of matrices until we arrive at the result. For example, to evaluate $\mathbf{E} = \mathbf{ABC}$ we can use our result above in which $\mathbf{AB} = \mathbf{D}$ and simply evaluate $\mathbf{E} = \mathbf{DC}$ using the technique previously mentioned.

---

Previously, we viewed the equation $\mathbf{Ax} = \mathbf{b}$ as representing a linear system of equations. Equivalently, we can view the result of $\mathbf{Ax}$ as a change of variables. In other words, we can write $\mathbf{y} = \mathbf{Ax}$ and interpret $\mathbf{y}$ as a new set of variables that are related to $\mathbf{x}$ through the transformation $\mathbf{A}$. In this course we will often pre-multiply a column vector $\mathbf{x}$ with $n$ elements (i.e., a $(n \times 1)$ matrix), by a matrix $\mathbf{A}$ of size $(m \times n)$ to produce this new vector $\mathbf{y}$ which has $m$ elements (i.e., a $(m \times 1)$ matrix). This equation is written

$$\mathbf{y} = \mathbf{Ax}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

The values of $\mathbf{y}$ can be evaluated using the same approach we described previously for multiplying matrices. For example, $y_1$ is obtained by multiplying the 1st row of $\mathbf{A}$ by the vector $\mathbf{x}$

$$y_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n$$

Repeating this for the elements $y_2, \cdots, y_m$ the vector $\mathbf{y}$ can be evaluated.

The *identity* matrix is a square $n$ by $n$ matrix, denoted $\mathbf{I}_n$ with all zero entries, except for ones on the diagonal.

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

An identity matrix has the property that when it pre-multiplies a vector $\mathbf{x}$ (with $n$ rows), or another matrix $\mathbf{A}$ (also with $n$ rows) they remain unchanged. In other other words,

$$\mathbf{I}_n\mathbf{A} = \mathbf{A}$$

$$\mathbf{I}_n\mathbf{x} = \mathbf{x}$$

---

**Example: Using the identity matrix**

For $n = 2$ the identity matrix is

$$\mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Then if $\mathbf{x} = (2 \quad 3)^{\mathrm{T}}$, we can show that

$$\mathbf{I}_2\mathbf{x} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1(2)+0(3) \\ 0(2)+1(3) \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \mathbf{x}$$

Similarly, if

$$\mathbf{A} = \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}$$

Then

$$\mathbf{I}_2\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1(3)+0(5) & 1(4)+0(6) \\ 0(3)+1(5) & 0(4)+1(6) \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \mathbf{A}$$

---

Now returning to the equation $\mathbf{y} = \mathbf{A}\mathbf{x}$, we often find ourselves in a situation where we have $\mathbf{A}$ and we have $\mathbf{y}$ but we must solve for $\mathbf{x}$. In scalar algebra, if we were presented with the equation $y = ax$ , we could easily solve for $x$ by dividing both sides by $1/a = a^{-1}$ to obtain $x = y/a$. Similarly, we define the *matrix inverse* $\mathbf{A}^{-1}$ with the $^{-1}$ superscript. $\mathbf{A}^{-1}$ has the property that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$. In other words, when a matrix is multiplied by its inverse, the product becomes the identity matrix. Thus when we multiply both sides of $\mathbf{y} = \mathbf{A}\mathbf{x}$ by the inverse $\mathbf{A}^{-1}$, we obtain the desired solution $\mathbf{x}$:

$$\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{x} \\ \mathbf{A}^{-1}\mathbf{y} &= \mathbf{A}^{-1}\mathbf{A}\mathbf{x} \\ \mathbf{A}^{-1}\mathbf{y} &= \mathbf{I}_n\mathbf{x} \\ \mathbf{A}^{-1}\mathbf{y} &= \mathbf{x} \end{aligned}$$

In MATLAB, the inverse function is implemented as `inv(A)`

---

**Example: Solve a LSEs using the matrix inverse**
Let

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 3 \\ -1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}$$

Then using MATLAB to compute the inverse

```
>> A = [1 0 -1;0  2  3;-1 1 0]
>> Ainv = inv(A)
Ainv =
   0.60000   0.20000  -0.40000
   0.60000   0.20000   0.60000
  -0.40000   0.20000  -0.40000
```

We can solve for $\mathbf{x}$ by computing $\mathbf{A}^{-1}\mathbf{y}$ as follows

```
>> x = Ainv*y
x =
   2.0000
   4.0000
  -1.0000
```

To double check that this is the correct $\mathbf{x}$ let's compute $\mathbf{Ax}$ and verify it is equal to $\mathbf{y}$

```
>> A*x
ans =
   3.0000
   5.0000
   2.0000
```

---

In general the inverse $\mathbf{A}^{-1}$ does not always exist and the LSEs can have: a) no solutions, b) one unique solution, or c) infinitely many solutions. In this course, we will assume $\mathbf{A}^{-1}$ is always well defined. Some additional useful properties for square matrices that we may come across are:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$
$$(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$$
$$(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}$$

## References

*Strang, G. (1988). Linear Algebra and Its Applications (Third Edition). Harcourt Brace Jovanovich.*