# Worksheet 3: Linear Combinations and Matrix-Vector Multiplication (§1.3)

**Problem 1: Column Vectors**. You are probably accustomed to representing points in the coordinate plane $\mathbb{R}^2$ or space $\mathbb{R}^3$ by "column vectors"—for example, $\begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}$ represents a particular point (on the $x$-axis) in 3-space. In Math 217, we think of points in $\mathbb{R}^n$ as "column vectors" (often shortened to "vector" in $\mathbb{R}^n$), where $n$ can be any natural number.

(a) Any two column vectors *of the same size* can be **added together**. Demonstrate by computing $\vec{e}_1 + \vec{a}$ for the vectors $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ in $\mathbb{R}^2$. How is this related to adding "vectors" represented by arrows by placing them head-to-tail as in Math 215 (or a physics class)?

(b) Any vector can be multiplied by a *scalar* (By the way, what is a "scalar"?). Demonstrate by computing $\pi\vec{a}$ for the vector $\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ in $\mathbb{R}^2$ and the scalar $\pi \in \mathbb{R}$. How is this related to scaling "vectors" represented by arrows as in Math 215 (or a physics class)?

(c) With $\vec{e}_1$ and $\vec{a}$ as above, draw a sketch of $\mathbb{R}^2$, which indicates the vectors $\vec{e}_1, \vec{a}, 2\vec{a}, \vec{e}_1 + 2\vec{a}$. Then shade in the sets

$$A = \{c\vec{a} \mid c \in \mathbb{R}\} \quad \text{and} \quad B = \{\lambda_1\vec{e}_1 + \lambda_2\vec{a} \mid 0 \le \lambda_1 \le 1,\ 0 \le \lambda_2 \le 2\}.$$

(d) Eventually we will define vectors to be any kind of mathematical objects where it makes sense to "add" and "scalar multiply". Can you come up with at least three examples of different kinds of mathematical objects, other than column vectors, that be "added" and "multiplied by scalars"? [Hint: think about row vectors, matrices, functions, or other objects you have been working with some since high school. Be precise if there are restrictions on sizes!]

**Solution:**

(a) Two column vectors can be added together if and only if they have the same number of components, in which case addition is performed "component-wise," like this:

$$\vec{a} + \vec{b} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{bmatrix}.$$

In this example,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Representing the vector by an arrow puts the tail of the vector at the origin and the head of the arrow at the coordinates of the column vector. Adding such vectors head to tail produces the same result as adding the coordinates of their heads.

(b) Any column vector can be multiplied by any scalar, and the scalar multiplication is performed "component-wise," like this:

$$c\vec{a} = c\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ca_1 \\ \vdots \\ ca_n \end{bmatrix}.$$

(c) The set $A$ is a line with slope 2 going through the origin. The set $B$ is a tall, skinny parallelogram (including its interior points) with vertices $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

(d) An easy example is row vectors, as long as we are adding row vectors of the same size. Another is matrices, again, they should be the same size to add them. A more interesting example is "real-valued functions on the real line": we can add two functions in the usual way (for example, $f(x) = x^2$ can be added to $g(x) = \cos x$ to get $(f+g)(x) = x^2 + \cos x$, and scaling by real numbers too (e.g. $\pi$ times the function $f(x) = \sin x$ gives us the function $\pi f(x) = \pi \sin x$. Yet another is "polynomials functions on the real line"—observe that adding/scaling polynomials produces another polynomial.

**Definition.** If $\vec{v}_1, \dots, \vec{v}_k$ is a list of vectors in $\mathbb{R}^n$ and $c_1, \dots, c_k$ is a list of scalars, then the vector

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$$

is called a **linear combination of the vectors** $\vec{v}_1, \dots, \vec{v}_k$ **with *coefficients*** $c_1, \dots, c_k$**.**

**Problem 2. Linear Combinations and Matrix Multiplication**.

(a) Let $\vec{a} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$. Compute the **linear combinations** $2\vec{a} + 3\vec{b}$ and $5\vec{a} - 2\vec{b}$.

(b) Consider the $4 \times 2$ matrix $A = \begin{bmatrix} \vec{a} & \vec{b} \end{bmatrix}$ whose columns are $\vec{a}$ and $\vec{b}$ (with $\vec{a}$ and $\vec{b}$ as in (a)). Compute the matrix products $A\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $A\begin{bmatrix} 5 \\ -2 \end{bmatrix}$. What do you notice when you compare with your answers in (a)?

(c) Let $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ in $\mathbb{R}^2$. For scalars $\lambda$ and $\mu$, compute the linear combination $\lambda\vec{a} + \mu\vec{b}$. Find a $2 \times 2$ matrix $A$ so that the product $A\begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \lambda\vec{a} + \mu\vec{b}$.

(d) Suppose $\vec{v}_1, \dots, \vec{v}_d$ are column vectors in $\mathbb{R}^n$. Given scalars $\lambda_1, \dots, \lambda_d \in \mathbb{R}$, express the linear combination vector $\lambda_1\vec{v}_1 + \cdots + \lambda_d\vec{v}_d \in \mathbb{R}^n$ as a *product of two matrices.* What are the sizes of your matrices? Discuss how to think of this as "block matrix multiplication" of a $1 \times d$ matrix" (with $n \times 1$ matrices as its entries) times a $d \times 1$ matrix.

(e) Let $A$ be an $n \times 3$ matrix and let $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ be the standard unit vectors in $\mathbb{R}^3$. Compute the matrix product $A\vec{e}_1$ in several examples to discover and state a general principle. What about $A\vec{e}_2$ and $A\vec{e}_3$? Now discuss the following general principle: *The $j$-th column of an $n \times d$ matrix $A$ is equal to $A\vec{e}_j$ where $\vec{e}_j$ is the standard unit vector in* $\mathbb{R}^d$. Deduce this **unreasonably useful lemma** from the principle discovered in (d).

**Solution:** For both (a) and (b) we get the same results: $\begin{bmatrix} 6 \\ 2 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -7 \\ 10 \end{bmatrix}$.

For (c), the linear combination is $\lambda\vec{a} + \mu\vec{b} = \begin{bmatrix} \lambda a_1 + \mu b_1 \\ \lambda a_2 + \mu b_2 \end{bmatrix}$. This is the same as the matrix product $\begin{bmatrix} \vec{a} & \vec{b} \end{bmatrix} \begin{bmatrix} \lambda & \mu \end{bmatrix} = \begin{bmatrix} \lambda a_1 + \mu b_1 \\ \lambda a_2 + \mu b_2 \end{bmatrix}$. So we can take $A$ to be the $2 \times 2$ matrix with first column $\vec{a}$ and second column $\vec{b}$.

(d): Let $A$ be the matrix with $d$ columns $\vec{v}_1, \vec{v}_2, \cdots \vec{v}_d$. Since each $\vec{v}_j \in \mathbb{R}^n$, the matrix $A = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_d \end{bmatrix}$ is $n \times d$. Then the linear combination $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_d\vec{v}_d$ is equal to $A \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$. One way to prove this is to write out each $\vec{v}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$ and then compute the product $A \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$ and compare it to the linear combination $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_d\vec{v}_d$. Both are the column vector

$$\begin{bmatrix} a_{11}\lambda_1 + \cdots + a_{1d}\lambda_d \\ a_{21}\lambda_1 + \cdots + a_{2d}\lambda_d \\ \vdots \\ a_{n1}\lambda_1 + \cdots + a_{nd}\lambda_d \end{bmatrix}.$$

(e). The general principle is this: let $\vec{e}_j \in \mathbb{R}^d$ be the "standard unit vector" that has zeros in each spot except in the $j$-th spot, where it has 1. Then for any $n \times d$ matrix $A$, the $j$-th column of $A$ is equal to the matrix product $A\vec{e}_j$. This is a special case of (d), since if $A = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_d \end{bmatrix}$, then from (d), $A\vec{e}_j$ is the linear combination $0\,\vec{v}_1 + \cdots + 1\,\vec{v}_j + \ldots 0\,\vec{v}_d$, where all terms are zero except the $j$-th, which is $\vec{v}_j$. This is called the **Unreasonably Useful Lemma** in the handout "Theory of Linear Algebra;" see Lemma 2.5 on page 7.

**Problem 3. Matrix Arithmetic**

(a) Any two matrices *of the same size* can be **added** together. Demonstrate by adding two together two $2 \times 3$ matrices $A$ and $B$ of your choice. Also write out a general formula for the sum of two arbitrary $m \times n$ matrices.

(b) Any matrix can be multiplied by any scalar. Demonstrate how to multiply one of your matrices in (a) by a scalar of your choice. Also write out a general formula for scalar multiplication of an arbitrary $m \times n$ matrix.

(c) If $A$ is an $m \times n$ matrix, what is $-A$? Demonstrate with your matrices in (a). What is $A + (-A)$? What is $A - B$? Is it the same as $A + (-B)$?

(d) How can we view row and column vectors as special kinds of matrices? What are their sizes (or dimensions)?

(e) **Caution:** Vectors can not be multiplied together willy-nilly! However, we *can* multiply a *row vector* of length $n$ by a column vector of length $n$. Demonstrate with an example where $n = 3$. What does this have to do with dot product from Math 215?

(f) Remind yourself what the *transpose* of a matrix $A$ means. What is the transpose of a column vector? A row vector? What is the size of the transpose of an $m \times n$ matrix?

(g) Writing $A^\top$ for the transpose of matrix $A$, explain why it is true that for all column vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$,

$$\vec{v}^\top \vec{w} \;=\; \vec{w}^\top \vec{v}.$$

[Hint: Remember some properties of dot product from Math 215.]

**Solution:**

(a) Two matrices can be added together if and only if they have the same size, in which case addition is performed "entry-wise," like this:

$$A + B \;=\; \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

(b) Any matrix can be multiplied by any scalar, and the scalar multiplication is performed "entry-wise," like this:

$$cA \;=\; c \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{bmatrix}.$$

(c) If $A$ is an $m \times n$ matrix, then $-A$ is just $(-1)A$, i.e., the matrix $A$ scaled by $-1$. The matrix $-A$ is the "additive inverse" of $A$, since $A + (-A) = (-A) + A = O$ where $O$ is the zero matrix of the same size as $A$. Subtraction of matrices makes good sense (as long as the two matrices have the same size!)

(d) A column vector in $\mathbb{R}^n$ is an $n \times 1$ matrix, and a row vector is a $1 \times n$ matrix.

(e) We multiply as follows:

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

which can also be written $\sum_{i=1}^{n} a_i b_i$.

(f) The transpose of a column vector (of length $n$) is a row vector (of length $n$), and vice versa. In general, the transpose turns the columns of a matrix into its rows. The transpose of a $m \times n$ matrix has size $n \times m$.

(g) If $\vec{x}, \vec{y} \in \mathbb{R}^n$ then $\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$. Since the dot product is commutative (Math 215), we therefore have

$$\vec{v}^T \vec{w} \;=\; \vec{v} \cdot \vec{w} \;=\; \vec{w} \cdot \vec{v} \;=\; \vec{w}^T \vec{v}.$$

**Problem 4: Linear Systems and Matrix-Vector Products.**

A *linear equation in one variable* is an equation of the form

$$ax = b \tag{†}$$

where $a$ and $b$ are real numbers and $x$ is a variable. You have known how to solve such equations since grade school.

(a) Write the general form of one linear equation in *two* variables. $ax + by = c$

(b) Write the general form of one linear equation in $n$ variables. $a_1 x_1 + \cdots + a_n x_n = b$

(c) What type of object is a solution of a linear equation in $n$ variables? A column vector of length $n$; i.e., an element of $\mathbb{R}^n$.

(d) How many solutions can a linear equation in $n$ variables have? What is the geometric shape of the solution set of a linear equation in $n$ variables? How do your answers depend on $n$? It has zero or infinitely many solutions, unless $n = 1$, in which case it has zero or one solution. The solution set, if nonempty, will be an $(n-1)$-dimensional hyperplane in $\mathbb{R}^n$.

(e) Write the general form of a *system of $m$ linear equations in $n$ variables*. Think carefully about how you want to name all the scalars and variables appearing in your equations.

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

(f) Now, rewrite your answer to part (e), which should have looked rather complicated, using matrices and (column) vectors in such a way that it resembles the very simple equation (†)

above.

$$A\vec{x} = \vec{b}, \qquad \text{where} \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

(g) Express the following linear system as a matrix-vector product of the form $A\vec{x} = \vec{b}$.

$$\begin{array}{ccccccccc} 0x_1 & + & 1x_2 & + & 2x_3 & + & 3x_4 & = & 1 \\ 4x_1 & + & 5x_2 & + & 6x_3 & + & 7x_4 & = & 2 \\ 8x_1 & + & 9x_2 & + & 0x_3 & + & 1x_4 & = & 3 \end{array}.$$

Let $A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 0 & 1 \end{bmatrix}$, let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$, and let $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

**Problem 5.** Let $\vec{v}_1, \vec{v}_2, \cdots \vec{v}_d$ be in $\mathbb{R}^n$ and let $A$ be the $n \times d$ matrix whose $j$-th column is the $n \times 1$ column vector $\vec{v}_j$. Prove that a vector $\vec{b} \in \mathbb{R}^n$ is a linear combination of $\vec{v}_1, \vec{v}_2, \cdots \vec{v}_d$ if and only if the system of linear equations $A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} = \vec{b}$ is *consistent.*

**Solution:** We need to prove two statements:

1. If $A\vec{x} = \vec{b}$ is consistent, then $\vec{b}$ is a linear combination of $\vec{v}_1, \vec{v}_2, \cdots \vec{v}_d$.
2. If $\vec{b}$ is a linear combination of $\vec{v}_1, \vec{v}_2, \cdots \vec{v}_d$, then $A\vec{x} = \vec{b}$ is consistent.

For (1), suppose the system $A\vec{x} = \vec{b}$ is consistent. This means that $\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_d \end{bmatrix}$ is a solution. That is, $A \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_d \end{bmatrix} = \vec{b}$. So by Problem 2, we have $\lambda_1 \vec{v}_1 + \cdots + \lambda_d \vec{d} = \vec{b}$, so that $\vec{b}$ is a linear combination as needed.

Conversely, say that $\vec{b}$ is a linear combination of $\vec{v}_1, \vec{v}_2, \cdots \vec{v}_d$. This means that there exists $\lambda_1, \ldots, \lambda_d$ such that $\vec{b} = \lambda_1 \vec{v}_1 + \cdots + \lambda_d \vec{d}$. Again, using the ideas of problem 2, this means that $A \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_d \end{bmatrix} = \vec{b}$, so that $\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_d \end{bmatrix}$ is a solution to the system $A\vec{x} = \vec{b}$. Because the system has a solution, it is consistent.