# Linear Transformations and Matrix Representations

Samuel R. Buss - Spring 2003

Revision 1.1.
(Corrections appreciated!)

These notes review the topics I lectured on while covering sections 4.1, 4.2, and 5.1 of the textbook *Linear Algebra with Applications* by S. Leon for my Math 20F class at UCSD. These notes do not cover everything you need to know from these chapters, so you should read the chapters as well.

The material is in slightly different order than that presented in class.

## 1 The theoretical constructions

### 1.1 Linear transformations

**Definition** Let $V$ and $W$ be vector spaces and $f : V \to W$. The function $f$ is called a *linear transformation* provided the following condition holds: For all $\mathbf{x}, \mathbf{y} \in V$ and all scalars $\alpha, \beta$,

$$f(\alpha\mathbf{x} + \beta\mathbf{y}) \;=\; \alpha f(\mathbf{x}) + \beta f(\mathbf{y}). \tag{1}$$

A linear transformation may also be called a "linear function," a "linear map", a "linear operator," etc. All these terms have the same meaning.

**Theorem 1** *If $f$ is a linear function, then $f(\mathbf{0}) = \mathbf{0}$.*

**Proof:** Let $\mathbf{x}, \mathbf{y}$ be arbitrary elements of $V$ and use $\alpha = \beta = 0$. Then $0{\cdot}\mathbf{x} = 0{\cdot}\mathbf{y} = \mathbf{0}$ and since $0 \cdot f(\mathbf{x}) = 0 \cdot f(\mathbf{y}) = \mathbf{0}$, equation (1) tells us that $f(\mathbf{0}) \;=\; \mathbf{0}$. □

**Theorem 2** *A function $f : V \to W$ is linear if and only if the following two conditions hold:*

a. *For all $\mathbf{x}, \mathbf{y} \in V$, $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$.*

b. *For all $\mathbf{x} \in V$ and all $\alpha \in \mathbb{R}$, $f(\alpha\mathbf{x}) = \alpha f(\mathbf{x})$.*

**Proof:** First suppose that conditions a. and b. hold. Then,

$$\begin{aligned} f(\alpha\mathbf{x} + \beta\mathbf{y}) &= f(\alpha\mathbf{x}) + f(\beta\mathbf{y}) && \text{by condition a.} \\ &= \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) && \text{by condition b.} \end{aligned}$$

Thus $f$ is linear. We have proved that conditions a. and b. together imply that $f$ is linear.

It remains to prove that any linear transformation satisfies conditions a. and b. To prove a., use (1) with $\alpha = \beta = 1$. To prove b., use (1) with $\beta = 0$. □

**Theorem 3** *Let $f$ be a linear transformation. Then, for all $\alpha_i$, $\mathbf{v}_i$,*

$$f(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) \;=\; \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2) + \cdots + \alpha_n f(\mathbf{v}_n).$$

**Proof:** To prove this, use equation (1) repeatedly. (The additions may be associated arbitrarily of course.) □

## 1.2 Matrix representation of a linear transformation

A basic property of linear transformations is that they can be represented by a matrix.

**Definition** Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then, we say that a $m \times n$ matrix $A$ *represents* the linear transformation $f$ iff

$$f(\mathbf{x}) \;=\; A\mathbf{x}, \qquad\qquad \text{for all } \mathbf{x} \in \mathbb{R}^n. \tag{2}$$

Note that the definition makes sense: If $A$ is $m \times n$, then for any $\mathbf{x} \in \mathbb{R}^n$, we have $A\mathbf{x} \in \mathbb{R}^m$. Thus, the mapping $\mathbf{x} \mapsto A\mathbf{x}$ is indeed a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. Indeed, the mapping $\mathbf{x} \mapsto A\mathbf{x}$ is clearly a linear map (you should prove this!), so every $m \times n$ matrix $A$ represents a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Now, we show how to construct the matrix $A$ that represents a linear function $f$. The next theorem tell us this is always possible.

**Theorem 4** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let $\mathbf{a}_i = f(\mathbf{e}_i)$ for all $i$ ($\mathbf{e}_i$ is the $i$-th standard basis vector). Then the matrix $A = (\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n)$ represents $f$.*

In words, the theorem says that the matrix $A$ that represents $f$ is formed by letting the columns of $A$ equal the values of $f(\mathbf{e}_i)$.

**Proof:** To prove this, first note that $A\mathbf{e}_i = \mathbf{a}_i = f(\mathbf{e}_i)$. For general $\mathbf{x} \in \mathbb{R}^n$, we have

$$\mathbf{x} \;=\; x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n,$$

so, by Theorem 3,

$$f(\mathbf{x}) \;=\; x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) + \cdots + x_n f(\mathbf{e}_n).$$

But, we also have $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$, which is equal to the same thing. Thus $A$ represents $f$. $\square$

## 1.3 Dot product as matrix product

Let $\mathbf{x} = (x_1\ x_2\ \cdots\ x_n)^T$ and $\mathbf{y} = (y_1\ y_2\ \cdots\ y_n)^T$.

**Definition** The *scalar product* of $\mathbf{x}$ and $\mathbf{y}$ is denoted $\mathbf{x} \cdot \mathbf{y}$ and is defined to equal

$$\mathbf{x} \cdot \mathbf{y} \;=\; x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

The scalar product is sometimes called the "dot product" or the "inner product."

The column vectors $\mathbf{x}$ and $\mathbf{y}$ are both $n \times 1$ matrices. Their tranposes, $\mathbf{x}^T$ and $\mathbf{y}^T$, are $1 \times n$ matrices. Thus, if we form the matrix product $\mathbf{x}^T\mathbf{y}$, or the matrix product $\mathbf{y}^T\mathbf{x}$, the result is a $1 \times 1$ matrix, that is to say, a scalar. In fact, this is the same as the scalar product.

**Theorem 5** $\mathbf{x} \cdot \mathbf{y} \;=\; \mathbf{x}^T\mathbf{y} \;=\; \mathbf{y}^T\mathbf{x}$.

**Proof:** The theorem is proved by inspection. For instance,

$$\mathbf{x}^T\mathbf{y} \;=\; (x_1\ x_2\ \cdots\ x_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \;=\; x_1y_1 + x_2y_2 + \cdots + x_ny_n \;=\; \mathbf{x}\cdot\mathbf{y}.$$

By a similar calculation, $\mathbf{y}^T\mathbf{x}$ also equals the dot product. □

Recall that the magnitude of a vector $\mathbf{x}$ is defined to be the square root of the sum of the squares of the components of the vector. That is, if $\mathbf{x} = (x_1\ x_2\ \cdots\ x_n)$, then

$$||\mathbf{x}|| \;=\; \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

The magnitude of $\mathbf{x}$ is the usual Euclidean length of the vector.

It is easy to check that $\mathbf{x}\cdot\mathbf{x} = ||\mathbf{x}||^2$. Thus we have:

**Theorem 6** $||\mathbf{x}||^2 = \mathbf{x}\cdot\mathbf{x} = \mathbf{x}^T\mathbf{x}$.

A remark on notation: Some authors use the notation $x$ to denote the magnitude of the vector $\mathbf{x}$. (Note the difference in fonts.) We will never use this notation in our course, however.

# 2 Examples

We now do a series of examples of linear transformations. Our examples are mostly in $\mathbb{R}^2$ and $\mathbb{R}^3$. As usual, we use the notation $\mathbf{x} = (x_1\ x_2)^T$ for vectors in $\mathbb{R}^2$ and $\mathbf{x} = (x_1\ x_2\ x_3)^T$ for for vectors in $\mathbb{R}^3$.

## 2.1 Scaling by the factor three

Let $f(\mathbf{x}) = 3\mathbf{x}$ be a function with domain and range $\mathbb{R}^2$. This function is easily proved to be linear, and it performs a scaling (or "magnification") by a factor of three. It is represented by the $2\times 2$ matrix

$$\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}.$$

## 2.2 Reflection across the line $y = x$.

The function that maps $\mathbf{x}$ to its reflection (or "mirror image") across the line $y = x$ can be defined by

$$f(\mathbf{x}) \;=\; \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}.$$

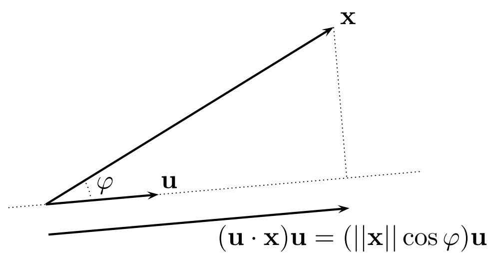


Figure 1: The projection of $\mathbf{x}$ onto the line parallel to a unit vector $\mathbf{u}$. $\varphi$ is the angle between $\mathbf{u}$ and $\mathbf{x}$. How would the picture look if $\varphi$ were between $90°$ and $180°$?

The matrix that represents this is $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. As expected, the first column of $A$ is $\binom{0}{1}$ which is the reflection of $\binom{1}{0}$ across the line $y = x$. Similarly, $A$'s second column is the reflection of $\binom{0}{1}$ across the same line.

### 2.3 Projection to the $x$-axis

The function that projects a point $\mathbf{x}$ orthogonally (perpendicularly) onto the $x$-axis is defined by $f(\mathbf{x}) = (x_1\ 0)^T$. Or, equivalently, this can be expressed as $f(\mathbf{x}) = x_1\mathbf{e}_1$.

The matrix that represents this function is $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

### 2.4 Projection onto the line $y = x$

The function $f$ that projects a point $\mathbf{x}$ onto the line $y = x$ can be defined by

$$f(\mathbf{x}) = \begin{pmatrix} \frac{1}{2}(x_1 + x_2) \\ \frac{1}{2}(x_1 + x_2) \end{pmatrix}.$$

One way to understand this definition of $f(\mathbf{x})$ is that it is equal to the average of $\mathbf{x}$ and its reflection across the line $y = x$ as defined two examples above.

The matrix $A$ that represents the projection $f$ is $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

### 2.5 Projection onto a line given by a unit vector $\mathbf{u}$

Let $\mathbf{u}$ be a unit vector, that is, $||\mathbf{u}|| = 1$. Let $L$ be the line through the origin that equals $Span(\mathbf{u})$. We shall define a function $f(\mathbf{x})$ that equals the projection of the point $\mathbf{x}$ onto the line $L$.

First, the so-called *scalar projection* of $\mathbf{x}$ onto the line $L$ is equal to $||\mathbf{x}|| \cos(\theta) = \mathbf{x} \cdot \mathbf{u}$. Here, $\theta$ is the angle between $\mathbf{x}$ and $\mathbf{u}$; see either Figure 1 above or the figure 5.1.2 on page 227 of the Leon textbook. The scalar projection is the length of the projection of $\mathbf{x}$ onto the line $L$.

Second, the so-called *vector projection* of $\mathbf{x}$ onto the line $L$ is equal to the scalar projection multiplied by $\mathbf{u}$, that is, it is equal to

$$(\mathbf{x} \cdot \mathbf{u})\mathbf{u}. \tag{3}$$

This vector projection is what is commonly called "the projection" of $\mathbf{x}$ onto $\mathbf{u}$.

To find a matrix that represents the function that projects $\mathbf{x}$ onto the unit vector $\mathbf{u}$, we rewrite the quantity (3) using matrix multiplications as

$$\begin{aligned}
(\mathbf{x} \cdot \mathbf{u})\mathbf{u} &= (\mathbf{u}^T\mathbf{x})\mathbf{u} \\
&= \mathbf{u}(\mathbf{u}^T\mathbf{x}) \qquad (\mathbf{u}^T\mathbf{x} \text{ is a scalar}) \\
&= (\mathbf{u}\mathbf{u}^T)\mathbf{x} \qquad (\text{associativity of matrix multiplication})
\end{aligned}$$

Thus, the matrix $\mathbf{u}\mathbf{u}^T$ represents the projection function $\mathbf{x} \mapsto (\mathbf{x} \cdot \mathbf{u})\mathbf{u}$.

What is the matrix $\mathbf{u}\mathbf{u}^T$? If we are working in $\mathbb{R}^n$, then $\mathbf{u}$ is an $n \times 1$ matrix and $\mathbf{u}$ is a $1 \times n$ matrix, so their product is a $n \times n$ matrix. For instance, in $\mathbb{R}^3$, it equals

$$\mathbf{u}\mathbf{u}^T = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} (u_1 \ u_2 \ u_3) = \begin{pmatrix} u_1^2 & u_1u_2 & u_1u_3 \\ u_1u_2 & u_2^2 & u_2u_3 \\ u_1u_3 & u_2u_3 & u_3^2 \end{pmatrix}.$$

This is an example of a symmetric matrix:

**Definition** A matrix $A$ is *symmetric* iff $A = A^T$.

## 2.6 Projection onto a line given by a general vector $\mathbf{u}$

The formulas in the previous example were correct only when $\mathbf{u}$ is unit vector (with magnitude equal to one). For general non-zero vectors $\mathbf{u}$, the same kind of analysis works: the only change is that you should first *normalize* $\mathbf{u}$ by dividing by its magnitude.

Thus, the vector projection of $\mathbf{x}$ onto the line through the origin in the direction $\mathbf{u}$ is equal to

$$\frac{\mathbf{x} \cdot \mathbf{u}}{||\mathbf{u}||^2}\mathbf{u} = \frac{\mathbf{u}^T\mathbf{x}}{\mathbf{u}^T\mathbf{u}}\mathbf{u}.$$

The matrix that represents this projection is the matrix

$$\frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}.$$

Note that this is also a symmetric matrix.

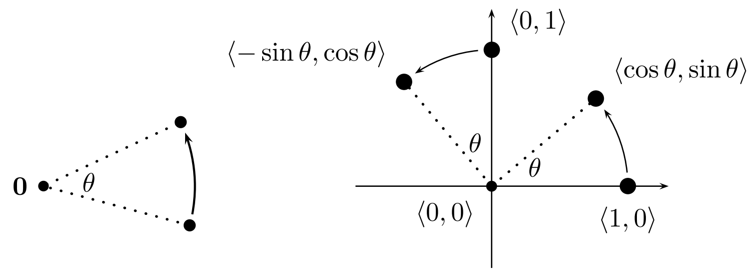


Figure 2: Effect of a rotation through angle $\theta$. The origin $\mathbf{0}$ is held fixed by the rotation.

### 2.7 Rotation in $\mathbb{R}^2$

Let $f(\mathbf{x})$ be the function that acts on the $xy$-plane as follows: the origin is kept fixed and all points in the plane are rotated through an angle of $\theta$ counterclockwise around the origin. For a picture of this, see Figure 2 above or the figure on page 200 of the textbook by Leon. This function can be seen to be linear. Thus, there is a $2 \times 2$ matrix that represents it. To find this matrix, note that

$$f : \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \qquad f : \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

Thus the linear transformation $f$ is represented by the matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

### 2.8 The cross-product as a linear transformation

Let $\mathbf{u}$ be a **fixed** vector in $\mathbb{R}^3$. Consider the function $f(\mathbf{x})$ defined by

$$f(\mathbf{x}) \;=\; \mathbf{u} \times \mathbf{x},$$

where $\times$ denotes the ordinary vector cross product in $\mathbb{R}^3$. Note that $f : \mathbb{R}^3 \to \mathbb{R}^3$. Also, it is easy to prove that $f$ is a linear transformation.

Therefore, it has a $3 \times 3$ matrix representation. To determine its matrix representation, note that, if $\mathbf{u} = (u_1\ u_2\ u_3)^T$, then

$$f : \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ u_3 \\ -u_2 \end{pmatrix} \qquad f : \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -u_3 \\ 0 \\ u_1 \end{pmatrix} \qquad f : \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} u_2 \\ -u_1 \\ 0 \end{pmatrix}.$$

Therefore, the matrix representation of the linear function $f$ is

$$A \;=\; \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}$$

You can confirm that $A$ is the correct matrix representation by checking that $A\mathbf{x} = \mathbf{u} \times \mathbf{x}$ holds for all $\mathbf{x}$.

Note that the matrix $A$ is an example of a skew-symmetric matrix:

**Definition** A matrix $A$ is *skew-symmetric* iff $A = -A^T$.

A skew-symmetric matrix must be square. Also, the diagonal elements of skew-symmetric matrix are equal to zero.