Math 130 Linear Algebra

Prof. D. Joyce, Clark University

Monday, 19 Oct 2009

Due today. Exercises from section 4.2: 8, 10ab, 11ab, 12ab, 14, 21ab, 23, ML2–ML5, ML8–ML9.

Due Wednesday. More exercises from section 4.2: 25, 26, 27ab, 31, 32, 34, T7, T10.

Quiz Friday on sections 4.1 and 4.2.

Due Monday. Exercises from section 4.3: 1, 4–7, 21, 22.

Due next Wednesday. More exercises from section 4.3: 25, 26, 27, 29, T9, T10, T11.

Last time. We extended our study of vectors, their lengths, and their inner products, to n-dimensional space. We showed that the triangle inequality

$$\|\mathbf{w} \pm \mathbf{v}\| \leq \|\mathbf{w}\| + \|\mathbf{v}\|$$

is equivalent to Cauchy's inequality

$$|\mathbf{w} \cdot \mathbf{v}| \leq \|\mathbf{w}\| \, \|\mathbf{v}\|,$$

and then proved Cauchy's inequality directly.

One consequence of Cauchy's inequality is that for any two vectors, $\mathbf{v}$ and $\mathbf{w}$, the value of

$$\frac{\mathbf{w} \cdot \mathbf{v}}{\|\mathbf{v}\| \, \|\mathbf{w}\|}$$

lies between -1 and 1. In the 2-dimensional case this equals $\cos\theta$ where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$. We can use that observation to define the angle θ between two vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbf{R}^n$ by

$$\theta = \arccos \frac{\mathbf{w} \cdot \mathbf{v}}{\|\mathbf{v}\| \, \|\mathbf{w}\|}.$$

With this definition angles between vectors have the properties we expect.

Today. Unit vectors in n-space, standard unit vectors, and linear transformations $L : \mathbf{R}^n \to \mathbf{R}^m$.

Unit vectors. A *unit vector* is a vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ whose length $\|\mathbf{u}\|$ equals 1. If a unit vector $\mathbf{u}$ is placed in standard position with its tail at the origin, then it's head will land at the point $(u_1, u_2, \ldots, u_n)$ on the unit $(n-1)$-sphere S^{n-1}. The $(n-1)$-sphere consists of points $(u_1, u_2, \ldots, u_n) \in \mathbf{R}^n$ such that

$$u_1^2 + u_2^2 + \cdots + u_n^2 = 1.$$

Note that the 2-sphere S^2 is what is ordinarily called the sphere in 3-space, while the 1-sphere S^1 is the ordinary unit circle in the plane. The higher dimensional spheres S^3, S^4, and so forth, are higher dimensional analogues.

We can interpret unit vectors as being directions in n-space.

Every vector $\mathbf{v}$ is the product of its length $\|\mathbf{v}\|$ and its direction $\mathbf{u}$ where $\mathbf{u}$ is the unit vector

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

so that $\mathbf{v} = \|\mathbf{v}\| \, \mathbf{u}$.

Standard unit vectors i, j, and k for 3-space $\mathbf{R}^3$. Our coordinatized 3-space $\mathbf{R}^3$ has three standard directions, the x-direction, the y-direction, and the z-direction. We can encode those three directions as unit vectors. We'll let $\mathbf{i}$ be the unit vector in the x-direction, namely, $\mathbf{i} = (1, 0, 0)$, $\mathbf{j}$ the unit vector in the y-direction, namely, $\mathbf{j} = (0, 1, 0)$, and $\mathbf{k}$ the unit vector in the z-direction, namely, $\mathbf{k} = (0, 0, 1)$

(We're using a bold face $\mathbf{i}$ to distinguish that direction from the italic i which is used in mathematics to denote the square root of -1.)

Every vector $\mathbf{v} = (x, y, z)$ can be uniquely written as a linear combination of these three standard unit vectors.

$$\mathbf{v} = (x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

There is an analogue for n dimensions. There are n standard unit vectors, often denoted

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, \ldots, 0), \\ \mathbf{e}_2 &= (0, 1, \ldots, 0), \\ &\ldots \\ \mathbf{e}_n &= (0, 0, \ldots, 1). \end{aligned}$$

Sometimes it's nice to have a notation without the ellipsis $(\ldots)$, and the Dirichlet delta symbol helps here. Let δ_{ij} be defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Then the j^{th} coordinate e_{ij} of the i^{th} standard unit vector is δ_{ij}.

Every vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ can be uniquely written as a linear combinations of these standard unit vectors

$$\mathbf{v} = (v_1, v_2, \ldots, v_n) = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n.$$

Definition of linear transformation $\mathbf{R}^n \to \mathbf{R}^m$. A linear transformation from one vector space to another is a function that preserves vector addition and scalar multiplication. In more detail, a *linear transformation* L from n-space to m-space is a function $L : \mathbf{R}^n \to \mathbf{R}^m$ such that

$$L(\mathbf{v} + \mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$$

and

$$L(c\mathbf{v}) = cL(\mathbf{v})$$

for every n-vector $\mathbf{v}$ and $\mathbf{w}$ and every scalar c.

Very often we are interested in the case when $m = n$.

We've already seen that a $m \times n$ matrix determines a linear transformation $\mathbf{R}^n \to \mathbf{R}^m$. If A is a $m \times n$ matrix, then $L(\mathbf{v}) = A\mathbf{v}$ is a linear transformation. Pretty soon, we'll be see that every linear transformation L is determined by such a matrix.

Back in chapter 1 we looked at some of these linear transformations of the plane. Things like reflections, rotations, dilations, contractions, and shears. There are others. For instance, if you compose a rotation and a contraction, you get a rotary contraction. The matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ describes a rotation about the origin by an angle of θ in the counterclockwise direction, while the matrix $\begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$ contracts the plane toward the origin by $\frac{1}{2}$, so the product of the matrices

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} = \begin{bmatrix} \frac{\cos\theta}{2} & -\frac{\sin\theta}{2} \\ \frac{\sin\theta}{2} & \frac{\cos\theta}{2} \end{bmatrix}$$

describes the compostion of first contracting then rotating. (In this case, the order doesn't matter since the two operations commute, but in many examples the order makes a difference.) The result is a kind of spiral transformation. If you take any point $\mathbf{x}$ and repeatedly apply this transformation, the orbit that is produced is a spiral toward the origin.

Properties of linear transformations. A few important properties follow directly from the definition. For instance, every linear transformation sends $\mathbf{0}$ to $\mathbf{0}$. Also, linear transformations preserve subtraction since subtraction can we written in terms of vector addition and scalar multiplication. A more general property is that linear transformations preserve linear combinations. For example, if $\mathbf{v}$ is a certain linear combination of other vectors $\mathbf{s}$, $\mathbf{t}$, and $\mathbf{u}$, say $\mathbf{v} = 3\mathbf{s}+5\mathbf{t}-2\mathbf{u}$, then $L(\mathbf{v})$ is the same linear combination of the images of those vectors, that is $L(\mathbf{v}) = 3L(\mathbf{s})+5L(\mathbf{t})-2L(\mathbf{u})$. This property can be stated as an identity

$$\begin{aligned} & L(c_1\mathbf{u_1} + c_2\mathbf{u_2} + \cdots + c_n\mathbf{u_n}) \\ = \; & c_1L(\mathbf{u_1}) + c_2L(\mathbf{u_2}) + \cdots + c_nL(\mathbf{u_n}) \end{aligned}$$

The property of preserving linear combinations is so important, that in some texts, a linear transformation is defined as a function that preserves linear combinations.

Linear transformations are always determined by matrices. Because linear transformations preserve linear combinations, we'll be able to find an $m \times n$ matrix A for any linear transformation $L : \mathbf{R}^n \to \mathbf{R}^m$. Here's how. Take a generic n-vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. Express $\mathbf{v}$ as a linear combination of the standard basis vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

as

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n.$$

Apply the linear transformation L which preserves this linear combination

$$\begin{aligned} L(\mathbf{v}) &= L(v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n) \\ &= v_1L(\mathbf{e}_1) + v_2L(\mathbf{e}_2) + \cdots + v_nL(\mathbf{e}_n) \end{aligned}$$

Name the m-vectors $L(\mathbf{e}_1)$ through $L(\mathbf{e}_n)$,

$$L(\mathbf{e}_1)= \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, L(\mathbf{e}_2)= \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, L(\mathbf{e}_n)= \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Then

$$\begin{aligned} L(\mathbf{v}) &= v_1L(\mathbf{e}_1) + v_2L(\mathbf{e}_2) + \cdots + v_nL(\mathbf{e}_n) \\ &= v_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + v_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} v_1a_{11} + v_2a_{12} + \cdots + v_na_{1n} \\ v_2a_{21} + v_2a_{22} + \cdots + v_na_{2n} \\ \vdots \\ v_1a_{m1} + v_2a_{m2} + \cdots + v_na_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = A\mathbf{v} \end{aligned}$$

Thus, we've seen that $L(\mathbf{v}) = A\mathbf{v}$, where the matrix A is the same for every vector v. In other words, we've represented an arbitrary linear transformation $L : \mathbf{R}^n \to \mathbf{R}^m$ as a matrix transformation $\mathbf{v} \mapsto A\mathbf{v}$. This matrix A, whose j^{th} column is the m-vector $L(\mathbf{e}_j)$, is sometimes called the *standard matrix* representing the linear transformation L.