# *Chapter 1: Functions on Euclidian Space*

*Notes on Spivak by Patch Kessler*

*December 29, 2018*

These notes and observations document my progression through Michael Spivak's Calculus on Manifolds[1]. The idea is to document my learning in beautiful Tufte-LaTeX style documents. Spivak's text is black, while all of my own writing is blue.

[1] Michael Spivak. *Calculus on Manifolds*. Benjamin Cummings, 1965. ISBN 0846590219

## *Norm and Inner Product*

Euclidean $n$-space $\mathbb{R}^n$ is defined as the set of all $n$-tuples $(x^1, ..., x^n)$ of real numbers $x^i$ (a "1-tuple of numbers" is just a number and $\mathbb{R}^1 = \mathbb{R}$, the set of all real numbers). An element of $\mathbb{R}^n$ is often called a point in $\mathbb{R}^n$, and $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$ are often called the line, the plane, and space respectively. If $x$ denotes an element of $\mathbb{R}^n$, then $x$ is an $n$-tuple of numbers, the i$^{th}$ one of which is denoted $x^i$; thus we can write

$$x = (x^1, ..., x^n).$$

A point in $\mathbb{R}^n$ is frequently also called a vector in $\mathbb{R}^n$, because $\mathbb{R}^n$, with $x + y = (x^1 + y^1, ..., x^n + y^n)$ and $ax = (ax^1, ..., ax^n)$, as operations, *is* a vector space (over the real numbers, of dimension $n$). In this vector space there is the notion of the length of a vector $x$, usually called the **norm** $|x|$ of $x$ and defined by $|x| = \sqrt{(x^1)^2 + \cdots + (x^n)^2}$. If $n = 1$, then $|x|$ is the usual absolute value of $x$. The relation between the norm and the vector space structure of $\mathbb{R}^n$ is very important.

I often use bold characters for vectors, and deal with them as columns.

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

This enables familiar mechanics and thought processes from linear algebra. I use subscripts rather than superscripts.

**1-1 Theorem.** If $x, y \in \mathbb{R}^n$ and $a \in \mathbb{R}$, then

(1) $|x| \geq 0$, and $|x| = 0$ if and only if $x = 0$.

(2) $|\sum_{i=1}^n x^i y^i| \leq |x| \cdot |y|$; equality holds if and only if $x$ and $y$ are linearly dependent.

(3) $|x + y| \leq |x| + |y|$.

(4) $|ax| = |a| \cdot |x|$.

**Proof**

(1) is left to the reader.

$|\mathbf{x}| \geq 0$
The real number $|x|$ is constructed by a process, the final step of which is a mapping by the square root function. The square root function returns values on $[0, \infty)$. □

$|\mathbf{x}| = 0$ if and only if $\mathbf{x} = \mathbf{0}$
If $\mathbf{x} = \mathbf{0}$, then each $x_i$ is zero, each $x_i^2$ is zero, the sum of these is zero, and the square root of this sum is zero. Conversely if $|\mathbf{x}| = 0$ then $x_1^2 + \cdots + x_n^2 = 0$ by the definition of $\sqrt{\ }$. The $x_i^2$ terms are non-negative, and so if any of them are positive, then the sum of them is positive. But the sum is not positive, and so none of the $x_i^2$ terms are positive. None of them are negative either, so they must all be zero. If $x_i^2$ is zero then $x_i$ is zero, and if all the $x_i$'s are zero then $\mathbf{x} = \mathbf{0}$. □

(2) If $x$ and $y$ are linearly dependent, equality clearly holds. If not, then $\lambda y - x \neq 0$ for all $\lambda \in \mathbb{R}$, so

$$\begin{aligned} 0 < |\lambda y - x|^2 &= \sum_{i=1}^n (\lambda y^i - x^i)^2 \\ &= \lambda^2 \sum_{i=1}^n (y^i)^2 - 2\lambda \sum_{i=1}^n x^i y^i + \sum_{i=1}^n (x^i)^2. \end{aligned}$$

Therefore the right side is a quadratic equation in $\lambda$ with no real solution, and its discriminant must be negative. Thus

$$4 \left( \sum_{i=1}^n x^i y^i \right)^2 - 4 \sum_{i=1}^n (x^i)^2 \cdot \sum_{i=1}^n (y^i)^2 < 0.$$

(3) $|x + y|^2 = \sum_{i=1}^n (x^i + y^i)^2$

$$\begin{aligned} &= \sum_{I=1}^n (x^i)^2 + \sum_{I=1}^n (y^i)^2 + 2 \sum_{I=1}^n x^i y^i \\ &\leq |x|^2 + |y|^2 + 2|x| \cdot |y| \quad \text{by (2)} \\ &= (|x| + |y|)^2. \end{aligned}$$

(4) $|ax| = \sqrt{\sum_{i=1}^n (ax^i)^2} = \sqrt{a^2 \sum_{i=1}^n (x^i)^2} = |a| \cdot |x|$. □

The quantity $\sum_{i=1}^n x^i y^i$ which appears in (2) is called the **inner product** of $x$ and $y$ and denoted $\langle x, y \rangle$. The most important properties of the inner product are the following.

**1-2 Theorem.** If $x$, $x_1$, $x_2$ and $y$, $y_1$, $y_2$ are vectors in $\mathbb{R}^n$ and $z \in \mathbb{R}$, then

(1) $\langle x, y \rangle = \langle y, x \rangle$ (symmetry).

(2) $\langle ax, y \rangle = \langle x, ay \rangle = a\langle x, y \rangle$ (bilinearity).
$\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$
$\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle$

(3) $\langle x, x \rangle \geq 0$, and $\langle x, x \rangle = 0$ if and only if $x = 0$ (positive definiteness).

(4) $|x| = \sqrt{\langle x, x \rangle}$.

(5) $4\langle x, y \rangle = |x + y|^2 - |x - y|^2$ (polarization identity).

**Proof**

(1) $\langle x, y \rangle = \sum_{i=1}^n x^i y^i = \sum_{i=1}^n y^i x^i = \langle y, x \rangle$.

(2) By (1) it suffices to prove

$$\langle ax, y \rangle = a\langle x, y \rangle,$$
$$\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle.$$

These follow from the equations

$$\begin{aligned} \langle ax, y \rangle &= \sum_{i=1}^n (ax^i)y^i = a \sum_{i=1}^n x^i y^i = a\langle x, y \rangle, \\ \langle x_1 + x_2, y \rangle &= \sum_{i=1}^n (x_1^i + x_2^i) y^i = \sum_{i=1}^n x_1^i y^i + \sum_{i=1}^n x_2^i y^i \\ &= \langle x_1, y \rangle + \langle x_2, y \rangle. \end{aligned}$$

(3) is left to the reader.

$\langle x, x \rangle$ is the sum of non-negative terms $x_i^2$, and so $\langle x, x \rangle \geq 0$. If $x = 0$, then from (1) of Theorem 1-1 we know that $|x| = 0$, and from (4) that $\langle x, x \rangle = 0$. Conversely if $\langle x, x \rangle = 0$, then $|x| = 0$ from (4), and from (1) of Theorem 1-1 we know that $x = 0$. □

(4) is left to the reader.
Expand $\langle x, x \rangle$ according to its definition, and take the square root, to get $\sqrt{x_1^2 + \cdots + x_n^2}$. This is the definition of $|x|$. □

(5) $|x+y|^2 - |x-y|^2$

$$\begin{aligned} &= \langle x+y, x+y\rangle - \langle x-y, x-y\rangle \quad \text{by (4)} \\ &= \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle - (\langle x,x\rangle - 2\langle x,y\rangle + \langle y,y\rangle) \\ &= 4\langle x,y\rangle. \ \square \end{aligned}$$

We conclude this section with some important remarks about notation. The vector $(0,...,0)$ will usually be denoted simply $0$. The **usual basis** of $\mathbb{R}^n$ is $e_1,...,e_n$, where $e_i = (0,...,1,...,0)$, with the 1 in the $i$th place. If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, the matrix of $T$ with respect to the usual bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ is the $m \times n$ matrix $A = (a_{ij})$, where $T(e_i) = \sum_{j=1}^m a_{ji}e_j$ - the coefficients of $T(e_i)$ appear in the $i$th column of the matrix. If $S : \mathbb{R}^m \longrightarrow \mathbb{R}^p$ has the $p \times m$ matrix $B$, then $S \circ T$ has the $p \times n$ matrix $BA$ [here $S \circ T(x) = S(T(x))$; most books on linear algebra denote $S \circ T$ simply $ST$]. To find $T(x)$ one computes the $m \times 1$ matrix

$$\begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} a_{11}, & \dots & , a_{1n} \\ \vdots & & \vdots \\ a_{m1}, & \dots & , a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix};$$

then $T(x) = (y^1,...,y^m)$. One notational convention greatly simplifies many formulas: if $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, then $(x,y)$ denotes

$$(x^1,...,x^n,y^1,...,y^m) \in \mathbb{R}^{n+m}.$$

Again, my thoughts are that it's simpler to deal with everything as columns in the first place. Also, for a clean and rigorous introduction to linear algebra, I recommend Linear Algebra Done Right[2].

[2] Sheldon Axler. *Linear Algebra Done Right*. Springer, 1996. ISBN 0387982582

**Problems**

**1-1.** Prove that $|x| \le \sum_{i=1}^n |x^i|$.

We can show this by writing $\mathbf{x}$ with respect to the $\mathbf{e}_i$ basis.

$$\mathbf{x} = x_1\mathbf{e}_1 + x_1\mathbf{e}_1 + ... + x_n\mathbf{e}_n$$

From Theorem 1-1 we know that $|x_i\mathbf{e}_i| = |x_i| \cdot |\mathbf{e}_i| = |x_i|$, and we also have the triangle inequality, which we now apply repeatedly

$$\begin{aligned} |\mathbf{x}| &= |x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + \cdots + x_n\mathbf{e}_n| \\ &\le |x_1| + |x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + \cdots + x_n\mathbf{e}_n| \\ &\le |x_1| + |x_2| + |x_3\mathbf{e}_3 + \cdots + x_n\mathbf{e}_n| \\ &\vdots \\ &\le |x_1| + |x_2| + |x_3| + \cdots + |x_n| \end{aligned}$$

**1-2.** When does equality hold in Theorem 1-1 (3)? *Hint:* Re-examine the proof; the answer is not "when $x$ and $y$ are linearly dependent."
When does $|\mathbf{x}+\mathbf{y}| = |\mathbf{x}| + |\mathbf{y}|$?

Start by writing $\mathbf{y}$ as $\alpha\mathbf{x} + \mathbf{e}$, where $\alpha \in \mathbb{R}$ and $\mathbf{e}$ is orthogonal to $\mathbf{x}$. Note that $|\mathbf{y}|^2 = \alpha^2|\mathbf{x}|^2 + |\mathbf{e}|^2$, and so

$$|\mathbf{y}| \geq \alpha|\mathbf{x}|, \tag{1}$$

with equality when $\mathbf{e} = \mathbf{0}$ and $\alpha > 0$. $\mathbf{e} = \mathbf{0}$ is another way of saying that $\mathbf{x}$ and $\mathbf{y}$ are parallel, while $\alpha > 0$ means that both vectors point in the same direction. Note that

$$|\mathbf{x}+\mathbf{y}|^2 = (1+\alpha^2)|\mathbf{x}|^2 + |\mathbf{e}|^2 + 2\alpha|\mathbf{x}|^2, \tag{2}$$

while

$$(|\mathbf{x}| + |\mathbf{y}|)^2 = (1+\alpha^2)|\mathbf{x}|^2 + |\mathbf{e}|^2 + 2|\mathbf{x}||\mathbf{y}|. \tag{3}$$

Combining (1), (2), and (3), we see that

$$|\mathbf{x}+\mathbf{y}|^2 \leq (|\mathbf{x}| + |\mathbf{y}|)^2, \tag{4}$$

with equality under the same conditions as in (1). If $a, b \geq 0$ and $a^2 \leq b^2$, then $a \leq b$, and so (4) implies the triangle inequality

$$|\mathbf{x}+\mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|. \tag{5}$$

We get equality under the same conditions as (1), namely when $\mathbf{x}$ and $\mathbf{y}$ are parallel and pointing in the same direction. This proof of the triangle inequality (i.e., (3) from Theorem 1-1) is slightly different from Spivak, making it easier to understand the conditions under which equality holds.

**1-3.** Prove that $|x-y| \leq |x| + |y|$. When does this equality hold?
This follows immediately from (5) in the previous exercise- simply replace $\mathbf{y}$ with $-\mathbf{y}$. We get equality when $\mathbf{x}$ and $\mathbf{y}$ are parallel and pointing in *opposite* directions.

**1-4.** Prove that $||x| - |y|| \leq |x-y|$.
From (2) in Theorem 1-1, we know that $|\langle \mathbf{x}, \mathbf{y}\rangle| \leq |\mathbf{x}||\mathbf{y}|$. From this we get $\langle \mathbf{x}, \mathbf{y}\rangle \leq |\mathbf{x}||\mathbf{y}|$, as well as $-|\mathbf{x}||\mathbf{y}| \leq -\langle \mathbf{x}, \mathbf{y}\rangle$. Using this, we see that

$$\begin{aligned}\langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle &= |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle \\ &\geq |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2|\mathbf{x}||\mathbf{y}| \\ &= (|\mathbf{x}| - |\mathbf{y}|)^2\end{aligned}$$

Thus $|\mathbf{x}-\mathbf{y}|^2 \geq (|\mathbf{x}| - |\mathbf{y}|)^2$, which implies $|\mathbf{x}-\mathbf{y}| \geq ||\mathbf{x}| - |\mathbf{y}||$.

**1-5.** The quantity $|y - x|$ is called the **distance** between $x$ and $y$. Prove and interpret geometrically the "triangle inequality": $|z - x| \leq |z - y| + |y - x|$.
From Problem 1-3, we know that $|\mathbf{u} - \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|$. Substituting in $\mathbf{u} = \mathbf{z} - \mathbf{y}$ and $\mathbf{v} = \mathbf{x} - \mathbf{y}$ gives the desired result. A geometrical interpretation is to think of $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ as the vertices of a triangle. The inequality establishes that the length of one side is always less than or equal to the sum of the lengths of the other two sides.

**1-6.** Let $f$ and $g$ be integrable on $[a, b]$.

(a) Prove that $|\int_a^b f \cdot g| \leq (\int_a^b f^2)^{\frac{1}{2}} \cdot (\int_a^b g^2)^{\frac{1}{2}}$. Hint: Consider separately the cases $0 = \int_a^b (f - \lambda g)^2$ for some $\lambda \in \mathbb{R}$ and $0 < \int_a^b (f - \lambda g)^2$ for all $\lambda \in \mathbb{R}$.

If we had established (2) in Theorem 1-1 for abstract vector spaces, we could prove this result by showing that the integrable functions on $[a, b]$ form a vector space, and that $\int_a^b f \cdot g$ is an inner product on this space. Instead of this, we stick with the scope of Spivak's development and follow his hint.

We wish to show

$$|\int_a^b fg| \leq (\int_a^b f^2)^{\frac{1}{2}} \cdot (\int_a^b g^2)^{\frac{1}{2}}. \tag{6}$$

Let $F(\lambda) = \int_a^b (f - \lambda g)^2$. Note that $F(\lambda) \geq 0$, because the integrand is never negative ($\bigstar$). Also, note that $F(\lambda) = A\lambda^2 + B\lambda + C$, where

$$A = \int_a^b g^2, \quad B = -2\int_a^b fg, \quad \text{and} \quad C = \int_a^b f^2.$$

If $A = 0$, then because of ($\bigstar$), we also have $B = 0$, and (6) is satisfied with both sides equal to zero.

If $A \neq 0$ and $F$ has no roots, then the discriminant $B^2 - 4AC$ is negative. Substituting in terms, we get

$$4(\int_a^b fg)^2 < 4(\int_a^b f^2)(\int_a^b g^2),$$

which leads to (6) as a strict inequality

If $A \neq 0$ and $F = 0$ for some $\lambda$, then from ($\bigstar$) we know that $F_{\text{min}} = 0$ as well. Because $A > 0$, $F_{\text{min}}$ is well defined, with value

$$F_{\text{min}} = F(-\frac{B}{2A}).$$

Setting this equal to zero, we obtain $B^2 = 4AC$, which leads to (6) as an equality.

(b) If equality holds, must $f = \lambda g$ for some $\lambda \in \mathbb{R}$? What if $f$ and $g$ are continuous?

From (a), we know that equality holds if $\int_a^b g^2 = 0$. Because of this, $f = \lambda g$ is **not** necessary for equality, even if $f$ and $g$ are continuous. For instance, let $f = \sin x$, $\lambda = 27$, and $g = 0$.

If we restrict ourselves to $\int_a^b g^2 > 0$, then from (a) we know that equality holds if $\int_a^b (f - \lambda g)^2 = 0$ for some $\lambda$.

These different cases can be combined into the single condition of *linear dependence*. For $f(x)$ and $g(x)$ to be linearly dependent, some linear combination of them must equal zero. That is, we must be able to find $\lambda_1, \lambda_2 \in \mathbb{R}$ such that

$$\lambda_1 f + \lambda_2 g = 0. \tag{7}$$

The condition we require is $\int_a^b (\lambda_1 f + \lambda_2 g)^2 = 0$. This is weaker than (7) because the integrand can be non zero at discrete points. This distinction goes away if $f$ and $g$ are continuous.

(c) Show that Theorem 1-1 (2) is a special case of (a).

Theorem 1-1 (2) is that $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq |\mathbf{u}||\mathbf{v}|$, with equality if and only if $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.

To get this from our work here, let $f$ and $g$ be piecewise constant, with values that equal the components of vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. For instance, let $f(x) = u_i$ and $g(x) = v_i$ for $x \in (i-1, i)$. This construction gives us

$$\langle \mathbf{u}, \mathbf{v} \rangle = \int_0^n f(x) \cdot g(x)\, dx.$$

The desired result $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq |\mathbf{u}||\mathbf{v}|$ then follows from (6). From (b), equality occurs when $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.

**1-7.** A linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is **norm preserving** if $|T(x)| = |x|$, and **inner product preserving** if $\langle Tx, Ty \rangle = \langle x, y \rangle$.

(a) Prove that $T$ is norm preserving if and only if $T$ is inner product preserving.

If $T$ is norm preserving then $T$ preserves inner products.

$$\begin{aligned}
\langle Tx, Ty \rangle &= \frac{1}{4}(|Tx + Ty|^2 - |Tx - Ty|^2) && \text{(polarization identity)} \\
&= \frac{1}{4}(|T(x+y)|^2 - |T(x-y)|^2) && \text{(linearity)} \\
&= \frac{1}{4}(|x+y|^2 - |x-y|^2) && \text{(norm preserving)} \\
&= \langle x, y \rangle && \text{(polarization identity)}
\end{aligned}$$

Conversely, if $T$ preserves inner products, then

$$|Tx| = \sqrt{\langle Tx, Ty\rangle} = \sqrt{\langle x, x\rangle} = |x|,$$

and so $T$ is norm preserving.

(b) Prove that such a linear transformation $T$ is 1-1 and $T^{-1}$ is of the same sort.

$T$ is 1-1 if $x \neq y$ implies $Tx \neq Ty$. This is equivalent to $Tx = Ty$ implying $x = y$, which we now show.

$$\begin{aligned} Tx = Ty &\Longrightarrow T(x-y) = 0 \\ &\Longrightarrow |T(x-y)| = 0 \\ &\Longrightarrow |x-y| = 0 \\ &\Longrightarrow x - y = 0 \\ &\Longrightarrow x = y \end{aligned}$$

We've used the linearity of $T$, the fact that $|\mathbf{u}| = 0$ if and only if $\mathbf{u} = \mathbf{0}$ (from Theorem 1-1), as well as the hypothesis that $T$ is norm preserving.

The inverse mapping $T^{-1}$ satisfies $T(T^{-1}x) = x$. As a result,

$$T^{-1}x = T^{-1}y \Longrightarrow T(T^{-1}x) = T(T^{-1}y) \Longrightarrow x = y,$$

and so $T^{-1}$ is 1-1 as desired. The question of whether $T^{-1}$ exists in the first place follows from linear algebra. If $T \in L(V, W)$, then

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

In our case, $\dim V = n$, and because $T$ is norm preserving, dim null $T = 0$. It follows that dim range $T = n$, and so $T$ is onto. We already know that $T$ is 1-1, and so it follows that $T$ is invertible.

**1-8.** If $x, y \in \mathbb{R}^n$ are non-zero, the **angle** between $x$ and $y$, denoted $\angle(x, y)$, is defined as $\arccos\left(\langle x, y\rangle / |x| \cdot |y|\right)$, which makes sense by Theorem 1-1 (2). The linear transformation $T$ is **angle preserving** if $T$ is 1-1, and for $x, y \neq 0$ we have $\angle(Tx, Ty) = \angle(x, y)$.

(a) Prove that if $T$ is norm preserving, then $T$ is angle preserving.

If $T$ is norm preserving, then from the previous problem $T$ is also inner product preserving. It follows that $T$ is angle preserving, because the angle is defined in terms of norms and inner products.

(b) If there is a basis $x_1, ..., x_n$ of $\mathbb{R}^n$ and numbers $\lambda_1, ..., \lambda_n$ such that $Tx_i = \lambda_i x_i$, prove that $T$ is angle preserving if and only if all $|\lambda_i|$ are equal.

Let $A(a, b)$ be the inner product of the unit vectors pointing in the directions of $a$ and $b$.

$$A(a, b) = \frac{\langle a, b\rangle}{|a||b|}.$$

Note that $\angle(a, b) = \arccos(A(a, b))$, and that $\angle(a, b) = \angle(c, d)$ if and only if $A(a, b) = A(c, d)$. From the polarization identity

$$4\langle x + y, x - y\rangle = |x|^2 - |y|^2,$$

we have

$$A(x + y, x - y) = \frac{|x|^2 - |y|^2}{4|x + y||x - y|}.$$

Let $e_i$ be the unit vector in the direction of each $x_i$, (i.e., $e_i = \frac{x_i}{|x_i|}$). Then $|e_i| = 1$, and so

$$A(e_i + e_j, e_i - e_j) = \frac{|e_i|^2 - |e_j|^2}{4|e_i + e_j||e_i - e_j|} = 0.$$

However, we also have $Te_i = \lambda_i e_i$, and so if $|\lambda_i| \neq |\lambda_j|$, then

$$A(T(e_i + e_j), T(e_i - e_j)) = \frac{|\lambda_i|^2 - |\lambda_j|^2}{4|\lambda_i e_i + \lambda_j e_j||\lambda_i e_i - \lambda_j e_j|} \neq 0.$$

Thus $|\lambda_i| \neq |\lambda_j|$ causes $T$ to not be angle preserving.

The other direction as stated is **false**. That is, all $|\lambda_i|$'s being equal does not make $T$ angle preserving. For instance, consider $\mathbb{R}^2$, with basis given by $x_1 = [1\ 0]^T$ and $x_2 = [1\ 1]^T$. Suppose $Tx_1 = x_1$, and $Tx_2 = -x_2$. It's easy to check that $\angle(x_1, x_2) = \pi/4$, while $\angle(Tx_1, Tx_2) = 3\pi/4$.

For the other direction to be true, we need all the $\lambda_i$ to be equal (not just their absolute values). If all the $\lambda_i$ are equal, say to $\lambda$, then for any vector $a \in \mathbb{R}^n$ we have $Ta = \lambda a$. As a result, for any two vectors $a, b \in \mathbb{R}^n$, we have

$$A(Ta, Tb) = \frac{\langle Ta, Tb\rangle}{|Ta||Tb|} = \frac{\langle \lambda a, \lambda b\rangle}{|\lambda a||\lambda b|} = \frac{\lambda^2}{|\lambda|^2} \cdot \frac{\langle a, b\rangle}{|a||b|} = A(a, b),$$

which establishes that $T$ is angle preserving.

(c) What are all angle preserving $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$?

$T$ is angle preserving if and only if there are orthonormal bases $\{e_i\}$ and $\{u_i\}$ of $\mathbb{R}^n$, and some positive $\lambda \in \mathbb{R}$ such that for all $i$

$$Te_i = \lambda u_i. \tag{8}$$

Let $T$ be angle preserving and let $\{e_i\}$ be any orthonormal basis of $\mathbb{R}^n$. Let $\lambda_i = |Te_i|$ and let $u_i$ be the unit vector in the direction of $Te_i$. All the $\lambda_i$'s are positive, and so we know from the same argument that we used in our answer to (b) that all the $\lambda_i$'s are equal. The orthogonality of the $u_i$'s is immediate, because if $u_i$ and $u_j$ are not orthogonal, then $A(u_i, u_j) \neq 0$. However $A(e_i, e_j) = 0$, and so this violates the hypothesis on $T$. Thus we started with an angle preserving $T$ and arrived at (8).

To show that (8) implies angle preservation by $T$, notice that $\langle Ta, Tb\rangle = \lambda^2\langle a, b\rangle$ for any two vectors $a, b \in \mathbb{R}^n$. Also, notice that $|Ta| = \lambda|a|$ and $|Tb| = \lambda|b|$. This gives us $A(Ta, Tb) = A(a, b)$, establishing that $T$ preserves angles.

**1-9.** If $0 \leq \theta < \pi$, let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ have the matrix $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. Show that $T$ is angle preserving and if $x \neq 0$, then $\angle(x, Tx) = \theta$.
$T$ satisfies the hypothesis of (c) from the previous problem, with $\{e_i\}$ equal to the standard basis, $u_1 = [\cos\theta \;\; -\sin\theta]^T$, $u_2 = [\sin\theta \;\; \cos\theta]^T$, and $\lambda = 1$.

$$T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix} \quad \text{and} \quad T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix}.$$

It follows from (c) of the previous problem that $T$ is angle preserving.

To show that $\angle(x, Tx) = \theta$ for nonzero $x = x_1e_1 + x_2e_2$, note that $|x|^2 = |Tx|^2 = x_1^2 + x_2^2$, and that $\langle x, Tx\rangle = (x_1^2 + x_2^2)\cos\theta$. Therefore

$$\frac{\langle x, Tx\rangle}{|x||Tx|} = \cos\theta.$$

From this, we get $\angle(x, Tx) = \theta$ as desired.

**1-10.**$^\star$ If $T : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ is a linear transformation, show that there is a number $M$ such that $|T(h)| \leq M|h|$ for $h \in \mathbb{R}^m$. Hint: Estimate $|T(h)|$ in terms of $|h|$ and the entries in the matrix of $T$.
We know from the singular value decomposition that associated with every $T : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ is an orthonormal basis $\{u_i\}$ of $\mathbb{R}^m$, an orthonormal basis $\{v_i\}$ of $\mathbb{R}^n$, and a collection of positive real *singular values* $\sigma_i$ such that $Tu_i = \sigma_i v_i$, for $i$ ranging from 1 to $\min(m, n)$. The $\sigma_i$ are the square roots of the eigenvalues of $TT^T$, and so come from a polynomial equation in the entries of the matrix of $T$. We can use the largest singular value $\sigma_{\max}$ as the $M$ requested in the problem statement.

We use the same symbol $T$ for the linear function and for the *matrix* of this linear function with respect to the standard basis.

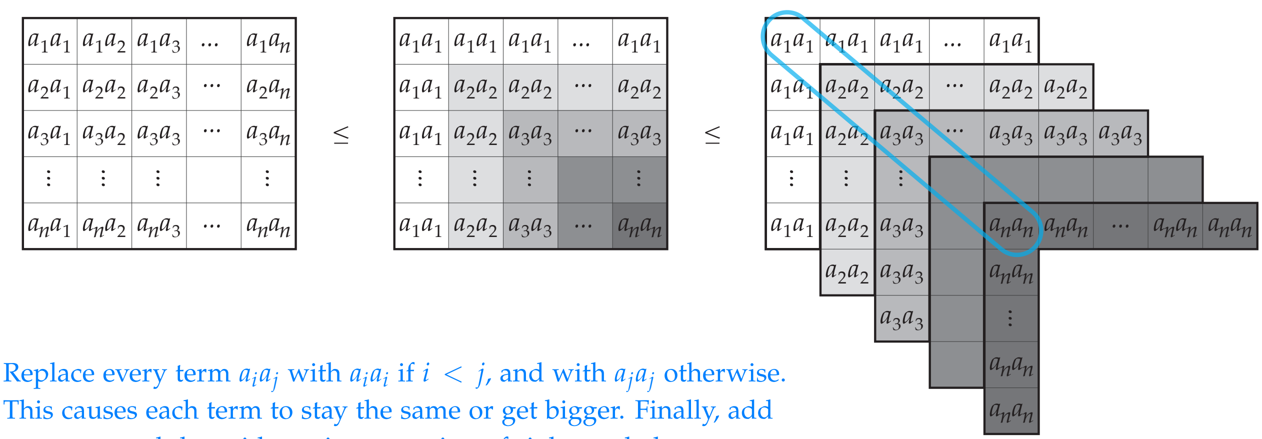
The singular value decomposition is like a Ferrari. It's a beautiful machine, and it'll get you to the corner store, however so will walking or riding a bike. There's an aesthetic appeal to tackling problems in the cleanest and simplest way possible, and so we try this again. The $i$th component of $Th$ is given by

$$g_i = T_{i1}h_1 + T_{i2}h_2 + \ldots + T_{im}h_m.$$

We'd like to place a bound on $\sqrt{g_1^2 + g_2^2 + \ldots + g_n^2}$. A key observation is that for any collection of non-negative $a_i$, we have

$$(a_1 + a_2 + \ldots + a_n)^2 \leq 2n(a_1^2 + a_2^2 + \ldots + a_n^2). \tag{9}$$

To see this, order and relabel the $a_i$'s so that $a_1 \geq a_2 \geq \ldots \geq a_n \geq 0$. Then write out the terms of $(a_1 + a_2 + \ldots + a_n)^2$ against a grid.




Replace every term $a_ia_j$ with $a_ia_i$ if $i < j$, and with $a_ja_j$ otherwise. This causes each term to stay the same or get bigger. Finally, add terms around the grid, getting a nesting of right angled arrangements of terms all with the same value. This results in $2n - 1$ diagonals, like the one outlined in blue, each of which sum to $a_1^2 + a_2^2 + \ldots + a_n^2$. We use (9) because it's true, and cleaner than including the $-1$. If $T_{\max}$ is the largest of the absolute values of the entries of $T$, then

$$\begin{aligned} g_i^2 &= (T_{i1}h_1 + T_{i2}h_2 + \ldots + T_{im}h_m)^2 \\ &\leq (|T_{i1}h_1| + |T_{i2}h_2| + \ldots + |T_{im}h_m|)^2 \\ &\leq T_{\max}^2(|h_1| + |h_2| + \ldots + |h_m|)^2 \\ &\leq T_{\max}^2 2m(h_1^2 + h_2^2 + \ldots + h_m^2) \\ &= T_{\max}^2 2m|h|^2 \end{aligned}$$

Summing all $n$ of the $g_i$'s, we see that $|Th|$ is bounded by $M|h|$ where

$$M = T_{\max}\sqrt{2mn}.$$

**1-11.** If $x, y \in \mathbb{R}^n$ and $z, w \in \mathbb{R}^m$, show that $\langle (x,z),(y,w)\rangle = \langle x,y\rangle + \langle z,w\rangle$ and $|(x,z)| = \sqrt{|x|^2 + |z|^2}$. Note that $(x,z)$ and $(y,w)$ denote points in $\mathbb{R}^{n+m}$.
We can see that $\langle (x,z),(y,w)\rangle = \langle x,y\rangle + \langle z,w\rangle$ by expanding $\langle (x,z),(y,w)\rangle$ according to the definition of the inner product.

$$\begin{aligned}\langle (x,z),(y,w)\rangle = \; & x_1y_1 + x_2y_2 + \ldots + x_ny_n \\ & + z_1w_1 + z_2w_2 + \ldots + z_mw_m\end{aligned}$$

The first and second rows on the right equal $\langle x,y\rangle$ and $\langle z,w\rangle$ respectively, again by the definition of the inner product. Using this result, observe that

$$\begin{aligned}|(x,z)| &= \sqrt{\langle (x,z),(x,z)\rangle} \\ &= \sqrt{\langle x,x\rangle + \langle z,z\rangle} \\ &= \sqrt{|x|^2 + |z|^2}\end{aligned}$$

**1-12.**$^\star$ Let $(\mathbb{R}^n)^\star$ denote the dual space of the vector space $\mathbb{R}^n$. If $x \in \mathbb{R}^n$, define $\varphi_x \in (\mathbb{R}^n)^\star$ by $\varphi_x(y) = \langle x,y\rangle$. Define $T : \mathbb{R}^n \longrightarrow (\mathbb{R}^n)^\star$ by $T(x) = \varphi_x$. Show that $T$ is a 1-1 linear transformation.
The dual space of $\mathbb{R}^n$ is the space of linear mappings from $\mathbb{R}^n$ to $\mathbb{R}$. In his prescription of $\varphi_x(y)$, Spivak constructs such mappings from vectors in $\mathbb{R}^n$. We need to show that this prescription is both linear and 1-1.

The linearity of $T$ follows from the linearity of the inner product used in the definition of $\varphi_x(y)$. If $a, b \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$, then

$$\begin{aligned}\varphi_{a+b}(y) &= \langle a+b, y\rangle \\ &= \langle a,y\rangle + \langle b,y\rangle \\ &= \varphi_a(y) + \varphi_b(y),\end{aligned}$$

and

$$\begin{aligned}\varphi_{\alpha a}(y) &= \langle \alpha a, y\rangle \\ &= \alpha\langle a,y\rangle \\ &= \alpha\varphi_a(y).\end{aligned}$$

For $T$ to be 1-1, we need to show that for any two unequal vectors $a, b \in \mathbb{R}^n$, we have $\varphi_a(x) \neq \varphi_b(x)$ for some $x \in \mathbb{R}^n$. Note that

$$\varphi_a(x) - \varphi_b(x) = \langle a,x\rangle - \langle b,x\rangle = \langle a-b, x\rangle$$

Picking $x = a - b$ causes $\varphi_a(x) - \varphi_b(x) = |a-b|^2$, which we know is non zero because $a \neq b$. (See (1) of Theorem 1-1.)

We might as well also show that $T$ is onto. Let $\varphi$ be any linear map from $\mathbb{R}^n$ to $\mathbb{R}$. Let $\{e_i\}$ be an orthonormal basis for $\mathbb{R}^n$, and note that for any vector $x = \sum_{i=1}^n x_i e_i$,

$$\varphi(x) = \sum_{i=1}^{n} x_i \varphi(e_i).$$

Define $a = \sum_{i=1}^n \varphi(e_i) e_i$, and note that

$$\langle a, x \rangle = \sum_{i=1}^{n} x_i \varphi(e_i).$$

It follows that $\varphi(x) = \langle a, x \rangle$, and so $T$ is onto.

**1-13.*** If $x, y \in \mathbb{R}^n$, then $x$ and $y$ are called **perpendicular** (or **orthogonal**) if $\langle x, y \rangle = 0$. If $x$ and $y$ are perpendicular, prove that $|x+y|^2 = |x|^2 + |y|^2$.
We are proving the Pythagorean Theorem in arbitrary dimensions!

$$\begin{aligned} |x+y|^2 &= \langle x+y, x+y \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle + 2\langle x, y \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle \\ &= |x|^2 + |y|^2 \end{aligned}$$

## *References*

Sheldon Axler. *Linear Algebra Done Right*. Springer, 1996. ISBN 0387982582.

Michael Spivak. *Calculus on Manifolds*. Benjamin Cummings, 1965. ISBN 0846590219.