# Module MAU23203: Analysis in Several Real Variables
# Michaelmas Term 2020
# Section 2: Schwarz's Inequality and some Related Inequalities

D. R. Wilkins



## Contents

# 2 Schwarz's Inequality and some Related Inequalities

## 2.1 Basic Properties of Vectors and Norms

We come now to a discussion of inequalities satisfied by vectors in the vector space $\mathbb{R}^n$ of dimension $n$ whose elements are ordered $n$-tuples of real numbers. The most basic inequality we consider here is *Schwarz's Inequality.* This inequality is then applied in the proof of the *Triangle Inequality.* It is further applied in the proof of some inequalities involving linear transformations and bilinear maps between finite-dimensional vector spaces.

The set $\mathbb{R}^n$ of ordered $n$-tuples of real numbers represents $n$-dimensional *Euclidean space* (with respect to the standard Cartesian coordinate system).

Let $\mathbf{x}$ and $\mathbf{y}$ be elements of $\mathbb{R}^n$, where

$$\mathbf{x} = (x_1, x_2, \ldots, x_n), \quad \mathbf{y} = (y_1, y_2, \ldots, y_n),$$

and let $c$ be a real number. We define

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \\ \mathbf{x} - \mathbf{y} &= (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n), \\ c\mathbf{x} &= (cx_1, cx_2, \ldots, cx_n), \\ \mathbf{x} \cdot \mathbf{y} &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n, \\ |\mathbf{x}| &= \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}. \end{aligned}$$

The quantity $\mathbf{x} \cdot \mathbf{y}$ is the *scalar product* (or *inner product*) of $\mathbf{x}$ and $\mathbf{y}$, and the quantity $|\mathbf{x}|$ is the *Euclidean norm* of $\mathbf{x}$. Note that $|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x}$. The *Euclidean distance* between two points $\mathbf{x}$ and $\mathbf{y}$ of $\mathbb{R}^n$ is defined to be the Euclidean norm $|\mathbf{y} - \mathbf{x}|$ of the vector $\mathbf{y} - \mathbf{x}$.

**Proposition 2.1** (Schwarz's Inequality) *Let $\mathbf{x}$ and $\mathbf{y}$ be elements of $\mathbb{R}^n$. Then $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$.*

**Proof** We note that $|t\mathbf{x} + \mathbf{y}|^2 \geq 0$ for all real numbers $t$. But

$$|t\mathbf{x} + \mathbf{y}|^2 = (t\mathbf{x} + \mathbf{y}).(t\mathbf{x} + \mathbf{y}) = t^2|\mathbf{x}|^2 + 2t\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2.$$

Therefore $t^2|\mathbf{x}|^2 + 2t\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2 \geq 0$ for all real numbers $t$. Thus $at^2 + bt + c \geq 0$ for all real numbers $t$, where $a = |\mathbf{x}|^2$, $b = 2\mathbf{x} \cdot \mathbf{y}$ and $c = |\mathbf{y}|^2$.

Now $at^2 + bt + c$ is a quadratic polynomial in the real variable $t$ whose values must be non-negative for all real values of $t$. A necessary and sufficient condition for this to be the case is that the inequality $b^2 \leq 4ac$ be satisfied. Thus, substituting in the values for $a$, $b$ and $c$ previously given, we find that

$$(\mathbf{x} \cdot \mathbf{y})^2 \leq |\mathbf{x}|^2 \, |\mathbf{y}|^2.$$

Schwarz's inequality now follows on taking the positive square roots of both sides. ∎

**Proposition 2.2** (Triangle Inequality) *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be elements of* $\mathbb{R}^n$*. Then* $|\mathbf{x}+\mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$.

**Proof** Using Schwarz's Inequality, we see that

$$\begin{aligned} |\mathbf{x}+\mathbf{y}|^2 &= (\mathbf{x}+\mathbf{y}).(\mathbf{x}+\mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\mathbf{x}\cdot\mathbf{y} \\ &\leq |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2|\mathbf{x}||\mathbf{y}| = (|\mathbf{x}| + |\mathbf{y}|)^2. \end{aligned}$$

The result follows directly. ∎

It follows immediately from the Triangle Inequality (Proposition 2.2) that

$$|\mathbf{z}-\mathbf{x}| \leq |\mathbf{z}-\mathbf{y}| + |\mathbf{y}-\mathbf{x}|$$

for all points $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ of $\mathbb{R}^n$. This important inequality expresses the geometric fact that the length of any one side of a triangle in a Euclidean space is less than or equal to the sum of the lengths of the other two sides of that triangle.

## 2.2 The Hilbert-Schmidt Norm of a Linear Transformation

Recall that the *length* (or *norm*) of an element $\mathbf{x} \in \mathbb{R}^n$ is defined such that

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

**Definition** Let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, and let $(T_{i,j})$ be the $n \times m$ matrix representing this linear transformation with respect to the standard bases of $\mathbb{R}^m$ and $\mathbb{R}^n$. The *Hilbert-Schmidt norm* $\|T\|_{\mathrm{HS}}$ of the linear transformation is then defined so that

$$\|T\|_{\mathrm{HS}} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} T_{i,j}^2}.$$

Note that the Hilbert-Schmidt norm is just the Euclidean norm on the real vector space of dimension $mn$ whose elements are $n \times m$ matrices representing linear transformations from $\mathbb{R}^m$ to $\mathbb{R}^n$ with respect to the standard bases of these vector spaces. Therefore it has the standard properties of the Euclidean norm. In particular it follows from the Triangle Inequality (Lemma 2.2) that

$$\|T + U\|_{\mathrm{HS}} \leq \|T\|_{\mathrm{HS}} + \|U\|_{\mathrm{HS}} \quad \text{and} \quad \|cT\|_{\mathrm{HS}} = |c|\, \|T\|_{\mathrm{HS}}$$

for all linear transformations $T$ and $U$ from $\mathbb{R}^m$ to $\mathbb{R}^n$ and for all real numbers $c$.

**Proposition 2.3** *Let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$. Then*

$$|T\mathbf{x}| \leq \|T\|_{\mathrm{HS}} |\mathbf{x}|$$

*for all $\mathbf{x} \in \mathbb{R}^m$, where $\|T\|_{HS}$ is the Hilbert-Schmidt norm of the linear transformation $T$.*

**Proof** Let $\mathbf{v} = T\mathbf{x}$, and let $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. Then

$$v_i = T_{i,1}x_1 + T_{i,2}x_2 + \cdots + T_{i,m}x_m$$

for all integers $i$ between 1 and $n$. It follows from Schwarz's Inequality (Lemma 2.1) that

$$v_i^2 \leq \left(\sum_{j=1}^{m} T_{i,j}^2\right)\left(\sum_{j=1}^{m} x_j^2\right) = \left(\sum_{j=1}^{m} T_{i,j}^2\right)|\mathbf{x}|^2.$$

Hence

$$|\mathbf{v}|^2 = \sum_{i=1}^{n} v_i^2 \leq \left(\sum_{i=1}^{n}\sum_{j=1}^{m} T_{i,j}^2\right)|\mathbf{x}|^2 = \|T\|_{\mathrm{HS}}^2 |\mathbf{x}|^2.$$

Thus $|T\mathbf{x}| \leq \|T\|_{\mathrm{HS}}|\mathbf{x}|$, as required. ∎

The following corollary follows immediately from Proposition 2.3.

**Corollary 2.4** *Let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$. Then there exists a positive real number $K$, dependent on the choice of linear transformation $T$ but independent of $\mathbf{x}$, with the property that*

$$|T\mathbf{x}| \leq K|\mathbf{x}|$$

*for all $\mathbf{x} \in \mathbb{R}^m$.*

In certain proofs in real analysis involving linear transformations, it is sufficient for the purposes of the proof that there should exist some positive constant $K$ for which the inequality in the statement of Corollary 2.4 is satisfied. Nevertheless, if a more precise estimate of the value of such a constant $K$ were required, then Proposition 2.3 would provide more information.

**Lemma 2.5** *Let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ and let $S\colon \mathbb{R}^n \to \mathbb{R}^p$ be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^p$. Then the Hilbert-Schmidt norm of the composition of the linear transformations $T$ and $S$ satisfies the inequality $\|ST\|_{\mathrm{HS}} \leq \|S\|_{\mathrm{HS}}\, \|T\|_{\mathrm{HS}}$.*

**Proof** The composition $ST$ of the linear transformations is represented by the product of the corresponding matrices. Thus the component $(ST)_{k,j}$ in the $k$th row and the $j$th column of the $p \times m$ matrix representing the linear transformation $ST$ satisfies

$$(ST)_{k,j} = \sum_{i=1}^{n} S_{k,i} T_{i,j}.$$

where $S_{k,i}$ and $T_{i,j}$ denote the components in the relevant rows and columns of the matrices representing the linear transformations $S$ and $T$ respectively. It follows from Schwarz's Inequality (Lemma 2.1) that

$$(ST)_{k,j}^2 \leq \left( \sum_{i=1}^{n} S_{k,i}^2 \right) \left( \sum_{i=1}^{n} T_{i,j}^2 \right).$$

Summing over $k$, we find that

$$\sum_{k=1}^{p} (ST)_{k,j}^2 \leq \left( \sum_{k=1}^{p} \sum_{i=1}^{n} S_{k,i}^2 \right) \left( \sum_{i=1}^{n} T_{i,j}^2 \right) = \|S\|_{\mathrm{HS}}^2 \left( \sum_{i=1}^{n} T_{i,j}^2 \right).$$

Then summing over $j$, we find that

$$\begin{aligned} \|ST\|_{\mathrm{HS}}^2 &= \sum_{k=1}^{p} \sum_{j=1}^{m} (ST)_{k,j}^2 \leq \|S\|_{\mathrm{HS}}^2 \left( \sum_{i=1}^{n} \sum_{j=1}^{m} T_{i,j}^2 \right) \\ &\leq \|S\|_{\mathrm{HS}}^2 \|T\|_{\mathrm{HS}}^2. \end{aligned}$$

On taking square roots, we find that $\|ST\|_{\mathrm{HS}} \leq \|S\|_{\mathrm{HS}}\, \|T\|_{\mathrm{HS}}$, as required. ∎

**Definition** Let $m$, $n$ and $p$ be positive integers. A function which assigns to every $m$-dimensional vector $\mathbf{x}$ and every $n$-dimensional vector $\mathbf{y}$ a $p$-dimensional vector $B(\mathbf{x}, \mathbf{y})$ determined by $\mathbf{x}$ and $\mathbf{y}$ is said to be a *bilinear map* if

$$\begin{aligned} B(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) &= B(\mathbf{x}_1, \mathbf{y}) + B(\mathbf{x}_2, \mathbf{y}) \\ B(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) &= B(\mathbf{x}, \mathbf{y}_1) + B(\mathbf{x}, \mathbf{y}_2) \\ B(c\mathbf{x}, \mathbf{y}) &= B(\mathbf{x}, c\mathbf{y}) = c\, B(\mathbf{x}, \mathbf{y}) \end{aligned}$$

for all $m$-dimensional vectors $\mathbf{x}$, $\mathbf{x}_1$, $\mathbf{x}_2$ in $\mathbb{R}^m$, all $n$-dimensional vectors $\mathbf{y}, \mathbf{y}_1, \mathbf{y}_2$ in $\mathbb{R}^n$, and all real numbers $c$.

The operation of multiplication by scalars on the real vector space $\mathbb{R}^n$ of ordered $n$-tuples of real numbers provides an example of a bilinear map from $\mathbb{R} \times \mathbb{R}^n$ to $\mathbb{R}^n$.

The scalar product on $\mathbb{R}^n$ provides an example of a bilinear map from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$.

The vector product on $\mathbb{R}^3$ provides an example of a bilinear map from $\mathbb{R}^3 \times \mathbb{R}^3$ to $\mathbb{R}^3$.

Throughout the discussion which immediately follows, let $B$ be a bilinear map which, for all $m$-dimensional vectors $\mathbf{x} \in \mathbb{R}^m$ and $n$-dimensional vectors $\mathbf{y} \in \mathbb{R}^n$ determines a $p$-dimensional vector $B(\mathbf{x}, \mathbf{y})$. Also let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p$ be the standard basis of $\mathbb{R}^p$, let $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_m$ be the standard basis of $\mathbb{R}^m$, and let $\mathbf{e}''_1, \mathbf{e}''_2, \ldots, \mathbf{e}''_n$ be the standard basis of $\mathbb{R}^n$, where, for each integer $i$ between 1 and $p$, each integer $j$ between 1 and $m$, and each integer $k$ between 1 and $n$, the vector $\mathbf{e}_i$ is that vector in $\mathbb{R}^p$ whose $i$th component is equal to one and whose other components are all zero, the vector $\mathbf{e}'_j$ is that vector in $\mathbb{R}^m$ whose $j$th component is equal to one and whose other components are all zero, and the vector $\mathbf{e}''_k$ is that vector in $\mathbb{R}^n$ whose $k$th component is equal to one and whose other components are all zero. Let coefficients $b_{i,j,k}$ be determined for $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, m$ and $k = 1, 2, \ldots, n$ such that

$$B(\mathbf{e}'_j, \mathbf{e}''_k) = \sum_{i=1}^{p} b_{i,j,k} \mathbf{e}_i$$

for $j = 1, 2, \ldots, m$ and $k = 1, 2, \ldots, n$. Now let $\mathbf{x} \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$, and let $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. Then

$$\mathbf{x} = \sum_{j=1}^{m} x_j \mathbf{e}'_j \quad \text{and} \quad \mathbf{y} = \sum_{k=1}^{n} y_k \mathbf{e}''_k.$$

It follows from the bilinearity of the map $B$ that

$$B(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{p} z_i \mathbf{e}_i,$$

where

$$z_i = \sum_{j=1}^{m} \sum_{k=1}^{n} b_{i,j,k}\, x_j\, y_k$$

for $i = 1, 2, \ldots, p$

**Proposition 2.6** *Let $B$ be a bilinear map associating to each $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$ an element $B(\mathbf{x}, \mathbf{y})$ of $\mathbb{R}^p$. And let $b_{i,j,k}$ be the coefficients determined for $j = 1, 2, \ldots, m$, $k = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, p$ so that if $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ then $B(\mathbf{x}, \mathbf{y}) = (z_1, z_2, \ldots, z_n)$, where*

$$z_i = \sum_{j=1}^{m} \sum_{k=1}^{n} b_{i,j,k}\, x_j\, y_k$$

*for $l = 1, 2, \ldots, p$. Then the Euclidean norms of $\mathbf{x}$, $\mathbf{y}$ and $B(\mathbf{x}, \mathbf{y})$ satisfy $|B(\mathbf{x}, \mathbf{y})| \leq K\, |\mathbf{x}|\, |\mathbf{y}|$ for all $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$, where*

$$K = \sqrt{\sum_{i=1}^{p} \sum_{j=1}^{m} \sum_{k=1}^{n} b_{i,j,k}^2}.$$

**Proof** Applying Schwarz's Inequality to the sum over the index $j$, we find that, for each integer $i$ between 1 and $p$,

$$z_i^2 \leq \left( \sum_{j=1}^{m} \left( \sum_{k=1}^{n} b_{i,j,k} y_k \right)^2 \right) \left( \sum_{j=1}^{m} x_j^2 \right).$$

Moreover a further application of Schwarz's inequality establishes that

$$\left( \sum_{k=1}^{n} b_{i,j,k} y_k \right)^2 \leq \left( \sum_{k=1}^{n} b_{i,j,k}^2 \right) \left( \sum_{k=1}^{n} y_k^2 \right).$$

Putting these inequalities together, we find that

$$z_i^2 \leq \left( \sum_{j=1}^{m} \sum_{k=1}^{n} b_{i,j,k}^2 \right) \left( \sum_{j=1}^{m} x_j^2 \right) \left( \sum_{k=1}^{n} y_k^2 \right).$$

Summing over $i$, we now find that

$$\sum_{i=1}^{p} z_i^2 \leq \left( \sum_{i=1}^{p} \sum_{j=1}^{m} \sum_{k=1}^{n} b_{i,j,k}^2 \right) \left( \sum_{j=1}^{m} x_j^2 \right) \left( \sum_{k=1}^{n} y_k^2 \right).$$

Taking square roots and applying the definition of the Euclidean norm, we find that

$$|B(\mathbf{x}, \mathbf{y})| \leq K\, |\mathbf{x}|\, |\mathbf{y}|$$

where

$$K = \sqrt{\sum_{i=1}^{p} \sum_{j=1}^{m} \sum_{k=1}^{n} b_{i,j,k}^2},$$

as required. ∎

The following corollary follows immediately from Proposition 2.6.

**Corollary 2.7** *Let $B$ be a bilinear map associating to each $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$ an element $B(\mathbf{x}, \mathbf{y})$ of $\mathbb{R}^p$. Then there exists a positive real number $K$, dependent on the choice of bilinear map $B$ but independent of $\mathbf{x}$ and $\mathbf{y}$, with the property that $|B(\mathbf{x}, \mathbf{y})| \leq K\, |\mathbf{x}|\, |\mathbf{y}|$ for all $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$.*

In certain proofs in real analysis involving bilinear maps, it is sufficient for the purposes of the proof that there should exist some positive constant $K$ for which the inequality in the statement of Corollary 2.7 is satisfied. Nevertheless, if a more precise estimate of the value of such a constant $K$ were required, then Proposition 2.6 would provide more information.