Orthogonality and Proportional Norms

By

P. Schöpf

durch das W. M. Ludwig Reich)

Abstract

Two norms on a real vector space define the same orthogonality relation
iff they are proportional. The aim of this note is to give a proof of this
statement with a minimum of results on convex sets, convex functions
and real analysis. Needed is only the right derivative of a convex function
and the théorème des accroissements finis as it is called by H. Cartan.

G. D. Birkhoff, R. C. James and others (see [1]) used in several important
papers the following concept of orthogonality in real normed linear
spaces $\langle X, \| \| \rangle$.

Definition 1. Let $x, y \in X$. We say $x \perp y$ ($x$ is orthogonal to $y$) iff
$\|x + \lambda y\| \geq \|x\|$ for all $\lambda \in \mathbb{R}$.

Let us now think of two norms $\| \|_1, \| \|_2$ on $X$, then we can ask, when
they determine the same orthogonality relation on $X$. If one analyzes the
geometrical meaning of this orthogonality relation, then it seems that the
following has to be true.

Theorem. Two norms $\| \|_1, \| \|_2$ on $X$ determine the same orthogonality
relation (i.e. $\perp_1 = \perp_2$) iff they are proportional (i.e. There exists a number $\sigma \in \mathbb{R}_{>0}$
with $\|x\|_1 = \sigma \|x\|_2$ for all $x \in X$.)

This theorem for example is useful in the proof of Theorem 4.17 in
[2] as Prof. R. Ger pointed out. Every proof of this theorem will use
several basic facts on convex sets and convex functions. Our aim is to give
a proof relying on a minimum of these facts and therefore we will use only, that a convex function has a right derivative at every inner point of its domain of definition. For this purpose let us fix some notations.

**Definition 2.** Let \((X, \|\|)\) be a real normed linear space and \(x, y \in X\) linearly independent. With \(x, y\) we always can define the following functions

\[
g(x, y; \bullet) : \mathbb{R} \to \mathbb{R}, \quad g(x, y; \mu) := \|x + \mu y\| \quad \text{for all } \mu \in \mathbb{R},
\]

\[
c(x, y; \bullet) : \mathbb{R} \to \mathbb{R}, \quad c(x, y; \mu) := \frac{1}{g(x, y; \mu)} (x + \mu y) \quad \text{for all } \mu \in \mathbb{R}.
\]

\(g^+(x, y; \bullet), c^+(x, y; \bullet)\) are the right derivatives of these functions.

**Lemma 1.** Let \((X, \|\|)\) be any real normed linear space, \(x, y \in X\) linearly independent, then the following statements are true.

1. \(c^+(x, y; 0) = -g^+(x, y; 0) \frac{x}{\|x\|^2} + \frac{y}{\|x\|}\)
2. \(c^+(x, y; 0)\) and \(c(x, y; 0)\) are linearly independent.
3. If \(x \perp y\), then \(-g^+(x, y; 0) \leq 0\).
4. If \(y := \eta \cdot x\) with \(\eta \in \mathbb{R}\), \(\eta > 0\), then
   \[
c^+(x, y; 0) = \eta c(x, y; 0)
\]
5. \(c(x, y - x; \lambda + \mu) = c(x + \lambda (y - x), y - x; \mu)\) and
   \[
c^+(x, y - x; \lambda) = c^+(x + \lambda (y - x), y - x; 0)
\]

**Proof:**

Ad 1. \(g^+(x, y; \mu)\) exists for every \(\mu \in \mathbb{R}\) because \(g(x, y; \bullet) : \mathbb{R} \to \mathbb{R}\) is a convex function. Differentiation therefore yields

\[
c^+(x, y; \mu) = -g^+(x, y; \mu) \frac{\partial g}{\partial \mu} (x + \mu y) = \frac{1}{g(x, y; \mu)} \frac{\partial}{\partial \mu} (x + \mu y)
\]

and with \(\mu = 0\) we get

\[
c^+(x, y; 0) = -g^+(x, y; 0) \frac{x}{g^2(x, y; 0)} + \frac{y}{g(x, y; 0)} = -g^+(x, y; 0) \frac{x}{\|x\|^2} + \frac{y}{\|x\|}.
\]

Ad 2. 1. shows that \(c^+(x, y; 0), x\) are linearly independent, because \(x, y\) are linearly independent. But \(x = \|x\| c(x, y; 0)\), hence \(c^+(x, y; 0), c(x, y; 0)\) are linearly independent.

Ad 3. \(x \perp y\) is defined by \(\|x + \mu y\| \geq \|x\|\) for all \(\mu \in \mathbb{R}\), but this implies that \(\mu = 0\) is an argument where the absolute minimum \(\|x\|\) of \(g(x, y; \bullet)\) is attained and therefore we must have \(-g^+(x, y; 0) \leq 0\).
Ad 4. One easily can check that
\[ c(x, y; \mu) = c\left(x, y; \frac{\mu}{\eta (1 - \kappa \mu)}\right) \]
for all \( \mu \in \mathbb{R} \) with \( \kappa \mu < 1 \). Differentiation yields
\[ c^+(x, y; \mu) = c^+\left(x, y; \frac{\mu}{\eta (1 - \kappa \mu)}\right) \frac{1}{\eta (1 - \kappa \mu)^2} \]
and for \( \mu = 0 \) we get the desired equation. □

Ad 5. A trivial computation.

Lemma 2. \( x \perp c^+(x, y; 0) \) for every pair of linearly independent vectors \( x, y \in X \).

Proof:
We only have to show that \( \| x + \lambda c^+(x, y; 0) \| \geq \| x \| \) for every \( \lambda \in \mathbb{R} \). For shorter notation we will write \( c(\delta) := c(x, y; \delta) \). Let \( \delta > 0 \) and \( \mu \notin [0, 1] \), we then get
\[ \| (1 - \mu) c(\delta) + \mu c(0) \| \geq (1 - \mu) \| c(\delta) \| + \mu \| c(0) \| = 1 = \| c(0) \| \]
or equivalently
\[ \left\| c(\delta) + (-\mu) \delta \left(\frac{c(\delta) - c(0)}{\delta}\right)\right\| \geq \| c(0) \|. \]
The last inequality says that for all \( \lambda \notin [-\delta, 0] \)
\[ \left\| c(\delta) + \lambda \left(\frac{c(\delta) - c(0)}{\delta}\right)\right\| \geq \| c(0) \|. \]
If we choose an \( \varepsilon > 0 \), then for every \( 0 < \delta < \varepsilon \) and every \( \lambda \notin [-\varepsilon, 0] \) we get
\[ \left\| c(\delta) + \lambda \left(\frac{c(\delta) - c(0)}{\delta}\right)\right\| \geq \| c(0) \|. \]
Taking the limit \( \delta \to 0 \) yields
\[ \| c(0) + \lambda c^+(0) \| \geq \| c(0) \| \]
for every \( \lambda \notin [-\varepsilon, 0] \). But \( \varepsilon \) was arbitrary and therefore
\[ \| x + \lambda c^+(0) \| \geq \| x \| \]
for all \( \lambda \in \mathbb{R} \). □
Lemma 3. Let \( \| \|_1, \| \|_2 \) be two norms on \( X \) which determine the same orthogonality relation \( \perp \) on \( X \) (i.e. \( \perp_1 = \perp_2 =: \perp \)). Let \( g_i(x,y;\bullet), \epsilon_i(x,y;\bullet) \) be the functions with respect to \( \| \|_i, i = 1, 2 \), then

\[
\epsilon_2^+(x,y;0) = \frac{\|x\|_1}{\|x\|_2} \epsilon_1^+(x,y;0).
\]

Proof:
By Lemma 1 we have

\[
\epsilon_1^+(x,y;0) = -g_1^+(x,y;0) \frac{x}{\|x\|_1} + \frac{y}{\|x\|_1} \quad \text{and} \quad \epsilon_2^+(x,y;0) = -g_2^+(x,y;0) \frac{x}{\|x\|_2} + \frac{y}{\|x\|_2}
\]

Substituting

\[
\tilde{y} := \epsilon_1^+(x,y;0) = \frac{1}{\|x\|_1} \left( y + \frac{-g_1^+(x,y;0)}{\|x\|_1} x \right)
\]

in these two equations yields (according to Lemma 1.4)

\[
\epsilon_1^+(x,\tilde{y};0) = \frac{\tilde{y}}{\|x\|_1} \quad \text{and} \quad \epsilon_2^+(x,\tilde{y};0) = \frac{1}{\|x\|_1} \epsilon_2^+(x,y;0)
\]

From this we get (by Lemma 1.1)

\[
\frac{1}{\|x\|_1} \epsilon_2^+(x,y;0) = \epsilon_2^+(x,\tilde{y};0) = -g_2^+(x,\tilde{y};0) \frac{x}{\|x\|_2} + \frac{\tilde{y}}{\|x\|_2}
\]

\[
= -g_2^+(x,\tilde{y};0) \frac{x}{\|x\|_2} + \frac{1}{\|x\|_2} \epsilon_1^+(x,y;0).
\]

By Lemma 2 we have \( x \perp \tilde{y} \) and Lemma 1.3 yields therefore

\[
-g_2^+(x,\tilde{y};0) \leq 0.
\]

If we substitute \( \tilde{y} =: \epsilon_2^+(x,y;0) \) in our starting equations, we get the analogous equation

\[
\frac{1}{\|x\|_2} \epsilon_1^+(x,y;0) = -g_1^+(x,\tilde{y};0) \frac{x}{\|x\|_2} + \frac{1}{\|x\|_1} \epsilon_2^+(x,y;0),
\]

with the analogous statement that

\[
-g_1^+(x,\tilde{y};0) \leq 0.
\]
Adding these last two equations yields

\[-g_2^+ (x, \tilde{y}; 0) \frac{x}{\|x\|_2^2} + -g_1^+ (x, \tilde{y}; 0) \frac{x}{\|x\|_1^2} = 0,\]

which implies that both coefficients are zero, because they are both \(\leq 0\).

Now we are ready, because of this we have

\[\frac{1}{\|x\|_2} \epsilon_2^+ (x, y; 0) = \frac{1}{\|x\|_1} \epsilon_1^+ (x, y; 0),\]

what we wanted.

**Proof of the theorem:**

Let \(\|\cdot\|_1, \|\cdot\|_2\) be two norms on \(X\), which define the same orthogonality relation \(\perp\) on \(X\) and \(x, y \in X\) linearly independent, then we will show that

\[\frac{\|y\|_1}{\|y\|_2} = \frac{\|x\|_1}{\|x\|_2}.\]

Our two curves are related in the following form

\[\epsilon_2(x, y - x; \lambda) \epsilon_1(x, y - x; \lambda)\]

and therefore we get by differentiation

\[\epsilon_2^+ (x, y - x; \lambda) = \left(\frac{g_1}{g_2}\right)^+ (x, y - x; \lambda) \epsilon_1 (x, y - x; \lambda)\]

\[+ \left(\frac{g_1}{g_2}\right)(x, y - x; \lambda) \epsilon_1^+ (x, y - x; \lambda).\]

Lemma 1.5, Lemma 3 and again Lemma 1.5 yield

\[\epsilon_2^+ (x, y - x; \lambda) = \epsilon_2^+ (x + \lambda (y - x), y - x; 0)\]

\[= \frac{\|x + \lambda (y - x)\|_1}{\|x + \lambda (y - x)\|_2} \epsilon_1^+ (x + \lambda (y - x), y - x; 0)\]

\[= \frac{\|x + \lambda (y - x)\|_1}{\|x + \lambda (y - x)\|_2} \epsilon_1^+ (x, y - x; \lambda)\]

i.e. \(\epsilon_2^+ (x, y - x; \lambda), \epsilon_1^+ (x, y - x; \lambda)\) are linearly dependent. On the other side we have according to Lemma 1.2, that \(\epsilon_1 (x, y - x; \lambda), \epsilon_1^+ (x, y - x; \lambda)\)
are linearly independent and hence we conclude, that
\[
\begin{pmatrix} g_1 \\ g_2 \end{pmatrix}^+ (x, y - x; \lambda) = 0
\]
for all $\lambda \in \mathbb{R}$. From this we get (by the théorème des accroissements finis [4])
\[
\left| \frac{g_1(x, y - x; 1)}{g_2(x, y - x; 1)} - \frac{g_1(x, y - x; 0)}{g_2(x, y - x; 0)} \right| \leq 0
\]
or, what is the same
\[
\frac{\|y\|_1}{\|y\|_2} = \frac{\|x\|_1}{\|x\|_2}.
\]
This implies the existence of a number $\sigma \in \mathbb{R}$, such that for all $x \in X$
\[
\|x\|_1 = \sigma \|x\|_2.
\]
The implication from proportionality of the two norms on equality of there associated orthogonality relations is trivial.

References

Author’s address: P. Schöpf, Institut für Mathematik, Universität Graz, Heinrichstraße 36/III, A-8010 Graz, Austria.