Inner product spaces

Laurence Barker, Department of Mathematics, Bilkent University

version: 26 April 2024

In current form, these notes are for supplementary material for two courses, MATH 220 Linear Algebra and MATH 224 Linear Algebra 2.

1: Real inner product spaces

For vector spaces over arbitrary fields, there is no general notion of distance or angle. In this section, we shall be confining attention to real vector spaces, that is, vector spaces over the field of real numbers $\mathbb{R}$. Sometimes, not always, when dealing with real vector spaces, consideration of distances or angles may be useful or necessary. In that case, the real vector spaces involved have to be viewed as something more than just real vector spaces. They have to be equipped with some or other kind of further structure. In this section, we shall be considering real vector spaces equipped with a further structure called an inner product.

Let V be a real vector space, we mean, a vector space over the field of real numbers $\mathbb{R}$. We define a **bilinear form** on V to be a function $\mathbb{R} \leftarrow V \times V$, written $\langle x \,|\, y \rangle \leftarrow\!\shortmid (x, y)$, such that the following two conditions hold:

Left linearity: For all $y \in V$, the function $\mathbb{R} \leftarrow V$ given by $\langle x \,|\, y \rangle \leftarrow\!\shortmid x$ is a linear map.

Right linearity: For all $x \in V$, the function $\mathbb{R} \leftarrow V$ given by $\langle x \,|\, y \rangle \leftarrow\!\shortmid y$ is a linear map.

Explicitly, the conditions are that, for all $x, x', y, y' \in V$ and $\lambda, \lambda \in \mathbb{R}$, we have

$$\langle \lambda x + \lambda' x' \,|\, y \rangle = \lambda \langle x \,|\, y \rangle + \lambda' \langle x' \,|\, y \rangle \;, \qquad\qquad \langle x \,|\, \lambda y + \lambda' y' \rangle = \lambda \langle x \,|\, y \rangle + \lambda' \langle x \,|\, y' \rangle \;.$$

We shall be considering two further conditions:

• The bilinear form is said to be **symmetric** provided $\langle x \,|\, y \rangle = \langle y \,|\, x \rangle$ for all $x, y \in V$.

• The bilinear form is said to be **positive-definite** provided $\langle x \,|\, x \rangle > 0$ for all $x \in V$ with $x \neq 0$.

We define an **inner product** on V to be a positive-definite symmetric bilinear form on V. We define a **real inner product space** to be a real vector space equipped with an inner product.

Let V be a real inner product space. For $x \in V$, we define the **norm** of x to be the non-negative real number

$$\|x\| = \sqrt{\langle x \,|\, x \rangle} \;.$$

Recall that, given $\lambda \in \mathbb{R}$, the **modulus** of λ is defined to be the non-negative real number $|\lambda|$ such that $|\lambda| = \lambda$ when $\lambda \geq 0$ whereas $|\lambda| = -\lambda$ when $\lambda \leq 0$. Thus, for all $\lambda \in \mathbb{R}$, we have $|\lambda| = \sqrt{\lambda^2}$.

Theorem 1.1: (Cauchy–Schwarz Inequality.) *Given a real inner product space V and $x, y \in V$, then*

$$|\langle x \,|\, y \rangle| \leq \|x\|.\|y\|$$

with equality if and only if $x = 0$ or $y = 0$ or $\mathbb{R}x = \mathbb{R}y$.

Proof: We may assume that $y \neq 0$. Given $t \in \mathbb{R}$, then

$$at^2 + bt + c = \|x + ty\|^2 \geq 0$$

where $a = \|y\|^2$ and $b = 2|\langle x \,|\, y \langle|$ and $c = \|x\|^2$. Note that $a \neq 0$. The graph of the function $at^2 + bt + c \leftarrow\!\shortmid t$ is a parabola ranging in the upper half place, possibly touching the horizontal axis, but not below that axis. So the discriminant of the quadratic equation $at^2 + bt + c = 0$ must be non-positive, $b^2 - 4ac \leq 0$, in other words, $(b/2)^2 \leq ac$. Taking a square root, we obtain the asserted inequality.

The argument also shows that $x + ty = 0$ for some t if and only if $(b/2)^2 = ac$. The rider follows. $\square$

For nonzero x and y in V, we define the **angle** between x and y to be the real number θ in the interval $0 \leq \theta < \pi$ such that

$$\cos(\theta) = \langle x \,|\, y \rangle / \|x\|.\|y\| \ .$$

The Cauchy–Schwarz Inequality assures us that θ is well-defined.

Theorem 1.2: (Triangle Inequality, first version.) *Given a real inner product space and $u, v \in V$, then*

$$\|u + v\| \leq \|u\| + \|v\| \ .$$

Proof: Applying the previous theorem, $\|u+v\|^2 = \|u\|^2 + 2\langle u \,|\, v\rangle + \|v\|^2 \leq (\|u\| + \|v\|)^2$. $\square$

Theorem 1.3: (Triangle Inequality, second version.) *Given a real inner product space and $x, y, z \in V$, then*

$$\|x - z\| \leq \|x - y\| + \|y - z\| \ .$$

Proof: Putting $u = x - y$ and $v = y - z$, then $u + v = x - z$. Thus, we have reduced to the previous version of the theorem. $\square$

Conversely, Theorem 1.2 can be quickly deduced from Theorem 1.3 by putting $x = u$ and $y = 0$ and $z = -v$. So the two theorems are really just two different ways of expressing one and the same theorem.

For $x, y \in V$, we define

$$d(x, y) = \|x - y\| \ .$$

We call $d(x, y)$ the **distance** between x and y. The Triangle Inequality says that

$$d(x, z) \leq d(x, y) + d(y, z) \ .$$

Put roughly, without digressing into any precise definition of a triangle, the theorem says that, for any triangle with vertices in a real inner product space, the length of any one of the edges is at most the sum of the lengths of the other two edges.

(Which of the above two versions of the Triangle Inequalty should we memorize for the exam? If one is asking that question, then one is doing it wrong. The trick is to understand, not to memorize. Everything about the Triangle Inequality can be recovered from just understanding what it is saying about the three edges of a triangle. That way, one sees that Theorems 1.2 and 1.3 have the same content. In a similar way, everything about the Cauchy–Schwartz Inequality can be recovered from just understanding how it tells us that the angle between two non-zero vectors is well-defined.)

♣ ♡ ♢ ♠

For any positive integer n, we shall equip the real vector space $\mathbb{R}^n$ with an inner product called the dot product. Thus, $\mathbb{R}^n$ will become a real inner product space. This is an important example. Indeed, as we shall later explain, every finite-dimensional real inner product space is, in a sense which we shall make precise, a copy of $\mathbb{R}^n$ equipped with the dot product.

Recall, the **standard basis** for $\mathbb{R}^n$ is defined to be the basis $\{e_1, ..., e_n\}$ such that, given $x \in \mathbb{R}^n$ and writing $x = (x_1, ..., x_n)$ with each $x_s \in \mathbb{R}$, then $x = \sum_s x_s e_s$.

We define the **dot product** on $\mathbb{R}^n$ to be the inner product $\langle - | - \rangle$ such that, given $x, y \in \mathbb{R}^n$ and writing $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$, then

$$\langle x \,|\, y \rangle = \sum_{s \in [1,n]} x_s y_s \ .$$

Note that the norm of x is

$$\|x\| = \sqrt{x.x} = \sqrt{\sum\nolimits_s x_s^2} \ .$$

The distance between x and y is

$$d(x, y) = \sqrt{\sum\nolimits_s (x_s - y_s)^2} \ .$$

When we speak of the inner product space $\mathbb{R}^n$, it is to be understood that, unless otherwise sdtated, the inner product under consideration is the dot product.

Often, as an alternative notation, the dot product is expressed using a dot, $x.y$ written instead of $\langle x \,|\, y \rangle$. That alternative notation can potentially be ambiguous because, in many different contexts, a dot is sometime used just to separate expressions that are being multiplied together, as in $2.3 = 6$. Still, if no multiplication operation on $\mathbb{R}^n$ is under consideration then, for $x, y \in \mathbb{R}^n$, the notation $x.y$ is unambiguous.

It is worth restating, in the special case $V = \mathbb{R}^n$, the two theorems in the previous section. The proofs are by substituting $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ into the statements of those theorems.

Theorem 1.4: (Cauchy–Schwarz Inequality, classical version.) *Given real numbers* x_1, ..., x_n, y_1, ..., y_n, *then*

$$\left(\sum\nolimits_r x_r y_r\right)^2 \leq \left(\sum\nolimits_s x_s^2\right) \left(\sum\nolimits_t y_t^2\right) \ .$$

Proof: Substitute $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ into Theorem 1.1. $\square$

Theorem 1.5: (Triangle Inequality, classical version.) *Given real numbers* u_1, ..., u_n, v_1, ..., v_n, *then*

$$\sqrt{\sum\nolimits_r (u_r + v_r)^2} \leq \sqrt{\sum\nolimits_s u_s^2} + \sqrt{\sum\nolimits_t v_t^2} \ .$$

Proof: Substitute $u = (u_1, ..., u_n)$ and $v = (v_1, ..., v_n)$ into Theorem 1.2. $\square$

$\clubsuit \qquad \heartsuit \qquad \diamondsuit \qquad \spadesuit$

FISH Introduce $\ell^2(\mathbb{R})$.

2: Orthonormal bases for real inner product spaces

Often, a system of linear equations has to be laboriously solved in order to express a given vector as a linear combination of the elements of a given basis. We shall be discussing, for real inner product spaces, a special kind of basis with the virtue that the coefficients of the linear combination can be calculated in a different way.

♣ ♡ ♢ ♠

To avoid confusion over some fundamental notions that involve some subtleties in the case of infinite-dimensional vector spaces, let us make a careful review of some background. Even for real or complex vector spaces, infinite sums of vectors do not always make sense, indeed, they can only make sense under suitable topological constraints. Such matters lie within the realm of functional analysis, and they lie outside our scope.

Recall that, for any subset $\mathcal{S}$ of a vector space X over any field F, we define a **linear combination** of elements of $\mathcal{S}$ to be a sum having the form

$$\sum_{s \in S} \lambda_s s$$

where each $\lambda_s \in F$ and there are only finitely many s such that $\lambda_s \neq 0$. Thus, although $\mathcal{S}$ may be infinite, the sum makes sense because it can be viewed as the finite sum of the nonzero terms. When the condition

$$\sum_s \lambda_s s = 0$$

implies that each $\lambda_s = 0$, we call $\mathcal{S}$ **linearly independent**. The set of linear combinations of elements of $\mathcal{S}$, denoted span($\mathcal{S}$), is a subspace of X. When span($\mathcal{S}$) $= X$, we say call $\mathcal{S}$ a **spanning set** for X. When $\mathcal{S}$ is a linearly independent spanning set for X, we call $\mathcal{S}$ a **basis** for X.

In functional analysis and its applications, a basis for X, as defined above, is sometimes called an **algebraic basis**, in distinction from another concept, called a **topological basis**, for which infinite sums may be considered.

♣ ♡ ♢ ♠

Often, a system of linear equations has to be laboriously solved in order to express a given vector as a linear combination of the elements of a given basis. We shall be discussing a special kind of basis where the coefficients of the linear combination can be calculated in a different way.

A subset $\mathcal{U}$ of a real inner product space V is said to be **orthogonal** provided $\langle u \,|\, u' \rangle = 0$ for all $u, u' \in S$ with $u \neq u'$. Obviously, any orthogonal subset of V is linearly independent. When $\mathcal{U}$ is an orthogonal subset of V and also a basis for V, we call $\mathcal{U}$ an **orthogonal basis** for V.

Remark 2.1: *Let V be a real inner product space with orthogonal basis $\mathcal{U}$. Given $x \in V$, then*

$$x = \sum_{u \in \mathcal{U}} \frac{\langle u \,|\, x \rangle}{\|u\|^2}\, u \,.$$

Proof: Since $\mathcal{U}$ spans V, we can write $x = \sum_u x_u u$ with each $x_u \in \mathbb{R}$, only finitely many of the x_u being nonzero. The required conclusion follows upon evaluating $\langle u \,|\, x \rangle$. □

An element $v \in V$ is called a **unit vector** in V provided $\|v\| = 1$. Note that, for any nonzero vector v in V, the vector $v/\|v\|$ is a unit vector. In casual language, the replacement of v with $v/\|v\|$ is called **normalization**.

A subset $\mathcal{W}$ of V is called **orthonormal** provided $\mathcal{W}$ is orthogonal and every element of $\mathcal{W}$ is a unit vector. When $\mathcal{W}$ is an orthonormal subset of V and also a basis for V, we call $\mathcal{W}$ an **orthonormal basis** for V.

Note that any orthogonal basis $\mathcal{V}$ for V gives rise to an orthonormal basis $\mathcal{W}$ for V given by

$$\mathcal{W} = \{v/\|v\| : v \in \mathcal{V}\} .$$

Thus, $\mathcal{W}$ is obtained from $\mathcal{V}$ by normalizing the elements. The next remark is just a special case of the previous one.

Remark 2.2: *Let V be a real inner product space with orthonormal basis $\mathcal{W}$. Given $x \in V$, then*

$$x = \sum_{w \in \mathcal{W}} \langle w \,|\, x \rangle w .$$

♣ ♡ ♢ ♠

Now let V be a finite-dimensional real inner product space. Let $n = \dim(V)$. Specializing a definition above, an orthogonal basis for V is a basis having the form $\{u_1, ..., u_n\}$ where $\langle u_s \,|\, u_t \rangle = 0$ for all $s, t \in [1, n]$ with $s \neq t$. Recall, the **Kronecker delta symbol** $\delta_{s,t}$ is defined to be the real number

$$\delta_{s,t} = \begin{cases} 1 & \text{if } s = t, \\ 0 & \text{if } s \neq t. \end{cases}$$

An orthonormal basis for V is a basis having the form $\{w_1, ..., w_n\}$ where

$$\langle u_s \,|\, u_t \rangle = \delta_{s,t} .$$

As an example, the standard basis $\{e_1, ..., e_n\}$ of $\mathbb{R}^n$ satisfies

$$\langle e_s \,|\, e_t \rangle = \delta_{s,t}$$

for all $s, t \in [0, n]$. In other words, the standard basis of $\mathbb{R}^n$ is an orthonormal basis for $\mathbb{R}^n$.

For ease of reference, we write down the following specialization of Remarks 2.3 and 2.4.

Remark 2.3: *Let V be a finite-dimensional real inner product space and let $x \in V$. Then:*

(1) *For any orthogonal basis $\{u_1, ..., u_n\}$ of V, we have* $x = \displaystyle\sum_{s \in [1,n]} \frac{\langle u_s \,|\, x \rangle}{\|u_s\|^2} u_s$.

(2) *For any orthonormal basis $\{w_1, ..., w_n\}$ of V, we have* $x = \displaystyle\sum_{s \in [1,n]} \langle w_s \,|\, x \rangle w_s$.

The next result shows that, when using coordinates with respect to an orthonormal basis, the inner product behaves very much like the dot product.

Corollary 2.4: *Let $\{w_1, ...w_n\}$ be an orthonormal basis for an inner product space V. Let $x, y \in V$. For each $s \in [1, n]$, write $x_s = \langle w_s \,|\, x\rangle$ and $y_s = \langle w_s \,|\, y\rangle$. Then:*

$$\langle x \,|\, y\rangle = \sum_s x_s y_s \, .$$

In particular $\|x\| = \sqrt{\sum_s x_s^2}$. *Also,* $d(x, y) = \sqrt{\sum_s (x_s - y_s)^2}$.

♣ ♡ ◇ ♠

We now describe a process, called the **Gram–Schmidt Process**, for replacing any given basis $\mathcal{U}$ of a finite-dimensional real inner product space with an orthogonal basis $\mathcal{V}$ and thence, if desired, an orthonormal basis $\mathcal{W}$.

Let V be a finite-dimensional real inner product space. Let $\mathcal{U} = \{u_1, ..., u_n\}$ be any basis for V. We define a set $\mathcal{V} = \{v_1, ..., v_n\}$ recursively by the condition that $v_1 = u_1$ and

$$v_r = u_r - \frac{\langle v_1 \,|\, u_r\rangle}{\|v_1\|^2}\, v_1 - ... - \frac{\langle v_{r-1} \,|\, u_r\rangle}{\|v_{r-1}\|^2}\, v_{r-1}$$

for $2 \leq r \leq n$. We define a set $\mathcal{W} = \{w_1, ..., w_n\}$ where each w_r is the normalization of v_r. That is to say, $w_r = v_r / \|v_r\|$.

Proposition 2.5: *With the notation above, $\mathcal{V}$ is an orthogonal basis for V. Also, $\mathcal{W}$ is an orthonormal basis for V.*

Proof: For $2 \leq s \leq n$, if the vectors v_1, ... , v_{s-1} are mutually orthogonal, then a direct calculation shows that $\langle v_t \,|\, v_s\rangle = 0$ for all $1 \leq t < s$. An inductive argument now yields the conclusion that $\mathcal{V}$ is an orthogonal basis for V. It follows immediately that $\mathcal{W}$ is an orthonormal basis for V. □

For any $x \in V$ with $x \neq 0$, we define the **normalization** of x to be the vector $x/\|x\|$. Note that the normalization $x' = x/\|x\|$ has norm $\|x'\| = 1$. We define $\mathcal{W} = \{w_1, ..., w_n\}$ where each w_r is the normalization of v_r. Thus,

$$w_r = v_r / \|v_r\| \, .$$

The previous remark immediately implies the next one.

Proposition 2.6: *With the notation above, $\mathcal{W}$ is an orthonormal basis for V.*

The latest proposition immediately yields the next theorem.

Theorem 2.7: *Any finite-dimensional real inner product space has an orthonormal basis.*

In the previous section, we promised to explain how any finite-dimensional real inner product space can be viewed as a copy of $\mathbb{R}^n$. We now fulfill that promise.

Given real inner product spaces U and V, we define an **isometry** $\alpha : U \leftarrow V$ to be an isomorphism $U \leftarrow V$ such that

$$\langle \alpha(v) \,|\, \alpha(v')\rangle = \langle v \,|\, v'\rangle \, .$$

When there exists an isometry $U \leftarrow V$, we say that U is isometric to V.

Consider inner product spaces U, V, W. The identity operator on U is an isometry. So U is isometric to U. Inverses of isometries are isometries. So if U is isometric to V, then V

is isometric to U. Composites of isometries are isometries. So if U is isometric to V and if V is isometric to W, then U is isometric to W. Thus, isometry of real inner product spaces is a formal equivalence relation. Intuitively, when U is isometric to V, we understand that, as real inner product spaces, U and V have the same structure, in other words, U and V are copies of each other. The next result says that every finite-dimensional real inner product space is, in the above sense, a copy of $\mathbb{R}^n$ for some n.

Corollary 2.8: *Let V be a finite-dimensional real inner product space. Let $n = \dim(V)$. Then V is isometric to $\mathbb{R}^n$.*

Proof: By Theorem 2.7, V has an orthonormal basis $\mathcal{W} = \{w_1, ..., w_n\}$. Let $\{e_1, ..., e_n\}$ be the standard basis for $\mathbb{R}^n$. By Corollary 2.4, there is an isometry $U \leftarrow \mathbb{R}^n$ such that $w_s \leftarrow\!\shortmid e_s$. □

♣ ♡ ♢ ♠

The Gram–Schmidt process adapts in a straightforward way to the case of a real inner product space with an infinite enumerated basis $\mathcal{U} = \{u_1, u_2, ...\}$. Defining $\mathcal{V} = \{v_1, v_2, ...\}$ and $\mathcal{W}\{w_1, w_2, ...\}$ by the same formulas as before, the proof of Proposition 2.5 shows that $\mathcal{V}$ is an orthogonal set and $\mathcal{W}$ is an orthogonal basis. Noting that $\mathrm{span}\{u_1, ...u_n\} = \mathrm{span}\{v_1, ..., v_n\}$ for each n, we see that $\mathcal{V}$ is a basis for V, hence $\mathcal{W}$ is a basis. In conclusion, $\mathcal{V}$ is an orthogonal basis for V and $\mathcal{U}$ is an orthonormal basis for V.

Let us give another example of a real inner product space. Consider an infinite sequence a_0, a_1, ... of real numbers such that $a_n \neq 0$ for only finitely many natural numbers n. Let $f : \mathbb{R} \leftarrow \mathbb{R}$ be the function such that

$$f(t) = a_0 + a_1 t + a_2 t^2 + ... = \sum_{m \in \mathbb{N}} a_m t^m \ .$$

We call f the **real polynomial function** with **coefficients** a_0, a_1, ... Note that f determines the coefficients a_m, indeed,

$$a_m = \frac{1}{m!} f^{(m)}(0)$$

where $f^{(m)}$ denotes the m-th derivative of f. If $a_n \neq 0$ and $a_m = 0$ for all $m \geq n$, then we say that f has **degree** n. In that case,

$$f(t) = a_0 + a_1 t + ... + a_n t^n \ .$$

It is to be understood that if $f(t) = 0$, then f has degree -1.

Let $P(\mathbb{R})$ denote the set of real polynomial functions. We make $P(\mathbb{R})$ become a real vector space such that, given $f, g \in P(\mathbb{R})$ then $(f + g)(t) = f(t) + g(t)$ and, given $a \in \mathbb{R}$, then $(af)(t) = a(f(t)$. Thus, we can write $af(t)$ unambiguously. We make $P(\mathbb{R})$ become a real inner product space with inner product given by

$$\langle f \,|\, g \rangle = \int_{-1}^{1} f(t) g(t) \,\mathrm{d}t \ .$$

Let I_0, I_1, ... be the real polynomial functions such that

$$I_n(t) = t^n \ .$$

Since a real polynomial function determines its coefficients, the set $\{I_0, I_1, ...\}$ is a basis for $P(\mathbb{R})$. But

$$\langle I_m \mid I_n \rangle = \int_{-1}^{1} I_m(t) I_n(t)\, \mathrm{d}t = \int_{-1}^{1} t^{m+n}\, \mathrm{d}t$$

$$= \left. \frac{t^{m+n+1}}{m+n+1} \right|_{t=-1}^{t=1} = \begin{cases} 2/(m+n+1) & \text{if } m+n \text{ is even,} \\ 0 & \text{if } m+n \text{ is odd.} \end{cases}$$

Evidently, $\{I_0, I_1, ...\}$ is not an orthogonal basis for $P(\mathbb{R})$.

For any nonzero real numbers a_0, a_1, ..., the set $\{a_0 I_0, a_1 I_1, ...\}$ is a basis for $P(\mathbb{R})$. in particular, defining

$$K_n(t) = \frac{1}{2^n} \binom{2n}{n} t^n$$

then $\{k_0, K_1, ...\}$ is a basis for $P(\mathbb{R})$. The **Legendre polynomial functions** L_1, L_1, ... are defined to be the real polynomial functions such that $\{L_0, L_1, ...\}$ is the orthogonal basis for $P(\mathbb{R})$ obtained from $\{K_0, K_1, ...\}$ by the Gram–Schmidt process without any normalization. We shall be investigating the Legendre polynomial functions in some of the exercises.

3: Orthogonal complements for real inner product spaces

FISH.

Theorem 3.1: *Given a subspace U of a finite-dimensional real inner product space V, then $V = U \oplus U^{\perp}$.*

Proof: FISH.

The conclusion of the theorem can fail when we drop the hypothesis that V is finite-dimensional. For example, putting $V = \ell^2(\mathbb{R})$, let U be the subspace of V consisting of those sequences x such that $x_i \neq 0$ for only finitely many natural numbers i. Then $U \neq V$ and $U^{\perp} = \{0\}$, hence $U \oplus U^{\perp} \neq V$.

FISH.

Procedural exercises on Sections 1, 2, 3

1.A: Let $u_1 = (2, 1, 1)$ and $u_2 = (1, 3, 1)$ and $u_3 = (1, 1, 3)$. Let $\{w_1, w_2, w_3\}$ be the orthonormal basis for $\mathbb{R}^3$ obtained from $\{u_1, u_2, u_3\}$ by the Gram–Schmidt process. Evaluate w_1 and w_2 and w_3.

1.B: Directly from the definition of the Legendre polynomial functions L_n in Section 2, give explicit formulas for $L_0(t)$, $L_1(t)$, $L_2(t)$, $L_3(t)$.

1.C:

Theoretical exercises on Sections 1, 2, 3

1.Z: Let $\langle - | - \rangle$ be a bilinear form on a real vector space V. We say that $\langle - | - \rangle$ is **non-degenerate** provided the following two conditions hold:

Left non-degeneracy: all $x \in V$ with $x \neq 0$, there exists $y \in V$ satisfying $\langle x \,|\, y \rangle \neq 0$.

Right non-degeneracy: for all $y \in V$ with $y \neq 0$ there exists $x \in V$ satisfying $\langle x \,|\, y \rangle \neq 0$.

(a) Suppose V is finite-dimensional. Show that $\langle - | - \rangle$ is left non-degenerate if and only if $\langle - | - \rangle$ is right non-degenerate.

(b) Give an example of an infinite-dimensional V such that the conclusion of part (a) fails.

1.Y: Let $\langle - | - \rangle$ be a non-degenerate bilinear form on a real vector space V. Show that $\langle x \,|\, x \rangle \neq 0$ for some $x \in V$.

1.X: Let V be a finite-dimensional inner product space and $\alpha : V \to V$ a function such that $\alpha(x, y) = d(\alpha(x), \alpha(y))$ for all $x, y \in V$. Show that α is an isometry. (Warning: most of the work is in showing that α is a linear map.)

1.W: Let V be a real inner product space with subspaces X and Y such that $V = X \oplus Y$. Let $\mathcal{U} = \{u_1, u_2, ...\}$ be a finite or countably infinite basis for V such that, for each index s, either $u_s \in X$ or $u_s \in Y$. Let $\mathcal{V}$ be the orthogonal basis obtained from $\mathcal{U}$ by the Gram–Schmidt process. Show that, for each s, if $u_s \in X$ then $v_s \in X$ whereas if $u_s \in Y$ then $v_s \in Y$.

1.V: A function $f : \mathbb{R} \leftarrow \mathbb{R}$ is said to be **even** provided $f(-t) = f(t)$ for all $t \in \mathbb{R}$, **odd** provided $f(-t) = -f(t)$ for all t. Directly from the definition of the Legendre polynomial functions $L_n$in Section 2, show that if n is even then L_n is even whereas if n is odd then L_n is odd.

1.U: In this question, you may assume the equality

$$\frac{1+m+n}{2^{m+n}} \sum_{a \in [0,m], b \in [0,n]} \binom{m}{a}^2 \binom{n}{b}^2 \int_{-1}^{1} (t-1)^{m+n-a-b} (t+1)^{a+b} \, \mathrm{d}t = \delta_{m,n}$$

for all natural numbers m and n.

(a) Directly from the above equality and the definition of the functions L_n in Section 2, show that

$$L_n(t) = \frac{1}{2^n} \sum_{b \in [0,n]} \binom{n}{b}^2 (t-1)^{n-b} (t+1)^b \ .$$

Hint: Consider the coefficient of t^n in the equality $(t+t)^{2n} = (t+1)^n (t+1)^n$.

(b) Evaluate $\|L_n\|^2$.

(c) Let f be a real polynomial function of degree at most n. Show that

$$f = \sum_{m \in [0,n]} b_m L_m$$

for some real numbers b_0, ..., b_n. Express b_m in terms of f and L_m.

Comment: I do not know of any way of proving the baroque equality above using just the basic techniques of calculus. The equality arises, eventually, after first using some more sophisticated techniques to establish the equivalence of various characterizations of the Legendre functions.

1.S: Let U be a subspace of a real inner product space V. Suppose that, for all $v \in V$, there exists $\pi(v) \in U$ such that $d(\pi(v), v)$ is minimal, in other words, $d(\pi(v), v) \leq d(u, v)$ for all $u \in U$.

(a) Show that $V = U \oplus U^{\perp}$.

(b) Show that π is a linear map.

Solutions to procedural exercises on Section 1

1.A: We have $w_1 = u_1/\|u_1\| = \dfrac{1}{\sqrt{6}}(2,1,1)$. Let $v_2 = u_2 - (w_1.u_2)w_1$. Then

$$v_2 = (1,3,1) - \left(\frac{1}{\sqrt{6}}(2,1,1).(1,3,1)\right)\frac{1}{\sqrt{6}}(2,1,1) = (1,3,1) - (2,1,1) = (-1,2,0)\ .$$

We have $w_2 = u_2/\|u_2\| = \dfrac{1}{\sqrt{5}}(-1,2,0)$. Let $v_3 = u_3 - (w_1.u_3)w_1 - (w_2.u_3)w_2$. Then

$$\begin{aligned} v_3 &= (1,1,3) - \left(\frac{1}{\sqrt{6}}(2,1,1).(1,1,3)\right)\frac{1}{\sqrt{6}}(2,1,1) - \left(\frac{1}{\sqrt{5}}(-1,2,0).(1,1,3)\right)\frac{1}{\sqrt{5}}(-1,2,0) \\ &= (1,1,3) - (2,1,1) - \frac{1}{5}(-1,2,0) = \frac{1}{5}\big((-5,0,10) - (-1,2,0)\big) = \frac{2}{5}(-2,-1,5)\ . \end{aligned}$$

We have $w_3 = v_3/\|v_3\| = (-2,-1,5)/\|(-2,-1,5)\| = \dfrac{1}{\sqrt{30}}(-2,-1,5)$. In conclusion,

$$w_1 = \frac{1}{\sqrt{6}}(2,1,1)\ , \qquad w_2 = \frac{1}{\sqrt{5}}(-1,2,0)\ , \qquad w_3 = \frac{1}{\sqrt{30}}(-2,-1,5)\ .$$

1.B: We begin with a little observation that will simplify the calculations. Let $f \in P(\mathbb{R})$ and FISH.

foxcat

We have

$$K_0(t) = 1\ , \qquad K_1(t) = t\ , \qquad K_2(t) = 3t^2/2\ , \qquad K_3(t) = 5t^3/2$$

for all $t \in \mathbb{R}$. Also, $L_0 = K_0$. We are to calculate

$$L_n = K_n - \frac{\langle K_0 \,|\, K_n \rangle}{\|K_0\|^2}\, K_0 - \ldots - \frac{\langle K_{n-1} \,|\, K_n \rangle}{\|K_0\|^2}\, K_{n-1}\ .$$

Since

foxcat

Solutions to theoretical exercises on Section 1

1.$\mathcal{A}$:

1.$\mathcal{B}$: