Summary of week 6 (lectures 16, 17 and 18)

Every complex number α can be uniquely expressed in the form $\alpha = a + b\boldsymbol{i}$, where a, b are real and $\boldsymbol{i} = \sqrt{-1}$. The *complex conjugate* of α is then defined to be the complex number $a - b\boldsymbol{i}$. We write $\overline{\alpha}$ for the complex conjugate of α.

Definition. Let V be a vector space over $\mathbb{C}$. An *inner product* on V is a function

$$\begin{aligned} V \times V &\to \mathbb{C} \\ (u, v) &\mapsto \langle u, v\rangle \end{aligned}$$

satisfying the following axioms: for all u, v, $w \in V$ and λ, $\mu \in \mathbb{C}$,

i) (*a*) $\langle u, \lambda v + \mu w\rangle = \lambda\langle u, v\rangle + \mu\langle u, w\rangle$;
 (*b*) $\langle \lambda u + \mu v, w\rangle = \overline{\lambda}\langle u, w\rangle + \overline{\mu}\langle v, w\rangle$;

ii) $\langle u, v\rangle = \overline{\langle v, u\rangle}$;

iii) $\langle u, u\rangle \in \mathbb{R}$, and $\langle u, u\rangle > 0$ if $u \neq 0$.

Note that the familiar "greater than" and "less than" relations for real numbers do not apply to complex numbers†; so for the statement $\langle u, u\rangle > 0$ to be meaningful, it is necessary for $\langle u, u\rangle$ to be a real number. Nevertheless, the part of Axiom iii) that says $\langle u, u\rangle \in \mathbb{R}$ is redundant, because it is a consequence of Axiom ii). Indeed, putting $v = u$ in Axiom ii) gives $\langle u, u\rangle = \overline{\langle u, u\rangle}$, which immediately implies that $\langle u, u\rangle \in \mathbb{R}$.

If V is a vector space over $\mathbb{R}$ then an inner product on V is a function $V \times V \to \mathbb{R}$ satisfying i), ii) and iii) above for all u, v, $w \in V$ and λ, $\mu \in \mathbb{R}$. In this case, of course, the complex conjugate signs can be omitted, since $\alpha = \overline{\alpha}$ when $\alpha \in \mathbb{R}$.

A real vector space equipped with an inner product is called a *real inner product space* or a *Euclidean space*, and a complex vector space equipped with an inner product is called a *complex inner product space* or a *unitary space*.

If V is a complex inner product space and $u \in V$ is fixed, then it follows from i) (a) that the function $f_u: V \to \mathbb{C}$ defined by

$$f_u(v) = \langle u, v\rangle \qquad \text{(for all } v \in V\text{)}$$

is linear, and hence, by Proposition 3.12 of [VST], that $\langle u, \underset{\sim}{0}\rangle = 0$. (Note that here $\underset{\sim}{0}$ denotes the zero element of V and 0 the zero of $\mathbb{C}$.) Since $\overline{0} = 0$, it follows from ii) that $\langle \underset{\sim}{0}, u\rangle = 0$ also.

There is further redundancy in the axioms: it is clear that ii) and i) (a) together imply i) (b).

† If "greater than" were defined for complex numbers in such a way that $\alpha > 0$ and $\beta > 0$ imply $\alpha\beta > 0$, and every nonzero α satisfies either $\alpha > 0$ or $-\alpha > 0$, then it would follow that $\alpha^2 > 0$ for all $\alpha \neq 0$, and hence $\beta > 0$ for all nonzero $\beta \in \mathbb{C}$.

A function f from one complex vector space to another is said to be *semilinear* or *conjugate linear* if $f(u+v) = f(u) + f(v)$ (for all u and v in the domain of f) and $f(\lambda u) = \overline{\lambda} f(u)$ for all u in the domain of f and all $\lambda \in \mathbb{C}$. Axiom i) (b) says that if $v \in V$ is fixed then $u \mapsto \langle u, v\rangle$ is a semilinear function $V \to \mathbb{C}$, while i) (a) says that if $u \in V$ is fixed then $v \mapsto \langle u, v\rangle$ is a linear function $V \to \mathbb{C}$. The inner product function is thus linear in one of its two variables and semilinear in the other; such functions are said to be *sesquilinear*. Because the complex conjugates disappear for real inner product spaces, the inner product is linear in both variables, or *bilinear*, in this case.

Definition. Vectors u, v in an inner product space V are said to be *orthogonal* if $\langle u, v\rangle = 0$. A set S of vectors is said to be an *orthonormal set* if $\langle v, v\rangle = 1$ for all $v \in S$ and $\langle u, v\rangle = 0$ for all u, $v \in S$ with $u \neq v$.

Proposition 5.4 of [VST] was proved in lectures: it asserts that a sequence of vectors that are nonzero and pairwise orthogonal is necessarily linearly independent. Consequently, such a sequence of vectors must form a basis for the subspace it spans.

Definition. An *orthogonal basis* of an inner product space is a basis whose elements are pairwise orthogonal.

We also proved the following important lemma (5.5 of [VST]).

Lemma. *Let $(u_1, u_2, \ldots, u_n)$ be an orthogonal basis of a subspace U of the inner product space V, and let $v \in V$. Then there is a unique element $u \in U$ such that $\langle x, u\rangle = \langle x, v\rangle$ for all $x \in U$, and it is given by $u = \sum_{i=1}^{n} \big(\langle u_i, v\rangle / \langle u_i, u_i\rangle\big) u_i$.*

It is in fact true that every finite-dimensional inner product space has an orthogonal basis. Moreover, if the dimension is at least 1 then there will be infinitely many orthogonal bases.† It is a consequence of Lemma 5.5 that if $(u_1, u_2, \ldots, u_n)$ and $(w_1, w_2, \ldots, w_n)$ are orthogonal bases of the same subspace U of V then $\sum_{i=1}^{n} \big(\langle u_i, v\rangle / \langle u_i, u_i\rangle\big) u_i = \sum_{i=1}^{n} \big(\langle w_i, v\rangle / \langle w_i, w_i\rangle\big) w_i$ for all $v \in V$. This was verified in Lecture 18 for the following example:

$$v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in V = \mathbb{R}^3; \qquad U = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \;\middle|\; x,\, y \in \mathbb{R} \right\};$$

$$(u_1, u_2) = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right); \qquad (w_1, w_2) = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right).$$

Other orthogonal bases of U, such as $\left(\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \right)$, give the same answer too.

† A zero-dimensional space has just one element—its zero element—and just one basis: the empty set.

Given any matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ we may define a function

$$\begin{aligned} R^n \times \mathbb{R}^n &\to R \\ (u, v) &\mapsto \langle u, v \rangle \end{aligned}$$

by the rule that $\langle u, v \rangle = {}^{\mathrm{t}}uAv$ for all u, $v \in \mathbb{R}^n$. This is always bilinear, as follows easily from standard properties of addition, multiplication and scalar multiplication for matrices. Indeed, if λ, $\mu \in \mathbb{R}$ and u, v, $w \in \mathbb{R}^n$ are arbitrary then we have

$$\begin{aligned} \langle \lambda u + \mu v, w \rangle &= {}^{\mathrm{t}}(\lambda u + \mu v)Aw \\ &= (\lambda\, {}^{\mathrm{t}}u + \mu\, {}^{\mathrm{t}}v)Aw \\ &= \lambda\, {}^{\mathrm{t}}uAw + \mu\, {}^{\mathrm{t}}vAw \\ &= \lambda \langle u, w \rangle + \mu \langle v, w \rangle, \end{aligned}$$

and a similar proof shows that $\langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle$. If A is a symmetric matrix, so that ${}^{\mathrm{t}}A = A$, then the product $(u, v) \mapsto \langle u, v \rangle$ is symmetric, in the sense that $\langle u, v \rangle = \langle v, u \rangle$ for all u and v, since

$$\langle u, v \rangle = {}^{\mathrm{t}}\langle u, v \rangle = {}^{\mathrm{t}}(\,{}^{\mathrm{t}}uAv) = {}^{\mathrm{t}}vA\,{}^{\mathrm{t}}({}^{\mathrm{t}}u) = \langle v, u \rangle.$$

A symmetric matrix A is said to be *positive definite* if ${}^{\mathrm{t}}uAu > 0$ for all nonzero $u \in \mathbb{R}^n$. When this condition is satisfied, $\langle u, v \rangle = {}^{\mathrm{t}}uAv$ defines an inner product on $\mathbb{R}^n$.

If ${}^{\mathrm{t}}u = (x_1, x_2, \ldots, x_n)$ and ${}^{\mathrm{t}}v = (y_1, y_2, \ldots, y_n)$ then

$${}^{\mathrm{t}}uAv = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j.$$

Hence we find that

$${}^{\mathrm{t}}uAu = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} (a_{ij} + a_{ji}) x_i x_j,$$

and if A is symmetric this becomes

$${}^{\mathrm{t}}uAu = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} a_{ij} x_i x_j.$$

Such an expression is called a *quadratic form* in the variables x_1, x_2, $\ldots$, x_n over the field $\mathbb{R}$. It is possible to use the technique known as *completing the square* to write any such expression in the form

$$\sum_{i=1}^{n} \varepsilon_i (\lambda_{i1} x_1 + \lambda_{i2} x_2 + \cdots + \lambda_{in} x_n)^2$$

where the coeffіcents ε_i are all either 0, 1 or -1, and the matrix whose (i,j) entry is λ_{ij} is invertible. Although we have not yet proved this result in lectures, some examples were given to illustrate the technique, and there are some other examples in [VST]. The matrix A is positive definite if and only if all the coefficients ε_i are equal to 1.

By Lemma 5.5, if U is a subspace of the inner product space V such that U has a finite orthogonal basis, there is a function $P\colon V \to U$ such that $\langle x, P(v)\rangle = \langle x, v\rangle$ for all $v \in V$ and all $x \in U$. Equivalently, $\langle x, v - P(v)\rangle = 0$ for all $v \in V$ and $x \in U$. The function P is called the *orthogonal projection* of V onto U. It follows readily from the formula

$$P(v) = \sum_{i=1}^{n} \bigl(\langle u_i, v\rangle / \langle u_i, u_i\rangle\bigr) u_i$$

(where $(u_1,\, u_2,\, \ldots,\, u_n)$ is any orthogonal basis for U) that P is a linear map.

We can use orthogonal projections to show that every finite-dimensional inner product space has an orthogonal basis. More generally, suppose that V is an inner product space and

$$U_1 \subset U_2 \subset \cdots U_d$$

is an increasing sequence of subspaces, with $\dim U_i = i$ for all i; then it is possible to find elements $u_1,\, u_2,\, \ldots\, u_d$ such that $(u_1,\, u_2,\, \ldots\, u_r)$ is an orthogonal basis of U_r, for each $r \in \{1, 2, \ldots, d\}$.

The procedure is inductive. First, note that since the dimension of U_1 is 1, all the nonzero elements of U_1 are scalar multiples of each other. Choose u_1 to be any nonzero element of U_1. Now, proceeding inductively, suppose that $1 < r \leq d$ and that we have already found elements $u_1,\, u_2,\, \ldots,\, u_{r-1}$ that are pairwise orthogonal and comprise a basis for U_{r-1}. Since $\dim U_r = r > r-1 = \dim U_{r-1}$, we can certainly find an element $v \in U_r$ such that $v \notin U_{r-1}$. Choose any such v, and put $u_r = v - P(v)$, where P is the orthogonal projection from U_r to U_{r-1}. We know that this orthogonal projection exists since our inductive hypothesis has given us an orthogonal basis for U_{r-1}. Observe that $u_r \neq 0$, since $P(v) \in U_{r-1}$ and $v \notin U_{r-1}$. Now by the defining property of orthogonal projections we know that $\langle x, u_r\rangle = \langle x, v - P(v)\rangle = 0$ for all $x \in U_{r-1}$. Hence $\langle u_j, u_r\rangle = 0$ for each $j \in \{1, 2, \ldots, u_{r-1}\}$. Since also $u_1,\, u_2,\, \ldots,\, u_{r-1}$ are pairwise orthogonal, it follows that $u_1,\, u_2,\, \ldots,\, u_r$ are pairwise orthogonal. By Proposition 5.4 it follows that $u_1,\, u_2,\, \ldots,\, u_r$ are linearly independent. Since they all lie in the r-dimensional space U_r, they form a basis for this space. So $(u_1,\, u_2,\, \ldots,\, u_r)$ is an orthogonal basis of U_r, and this completes the inductive proof.

The *Gram-Schmidt orthogonalization process* is a procedure that starts with a linearly independent sequence of vectors $(v_1, v_2, \ldots, v_d)$ in an inner product space and produces an orthogonal sequence of vectors $(u_1, u_2, \ldots, u_d)$ such that $\mathrm{Span}(u_1, u_2, \ldots, u_r) = \mathrm{Span}(v_1, v_2, \ldots, v_r)$ for all r from 1 to d. The relevant formulas (which were given in lectures) can be found on p.108 of [VST].

Note in particular that if U is any finite-dimensional subspace of the inner product space V then an orthogonal basis for U can be found, and therefore the projection of V onto U exists.

In any inner product space V the *distance* between elements $u, v \in V$ is defined to be the real number $d(u,v) = \|u-v\|$. It follows from iii) of the definition of inner product that $d(u,v) > 0$ if $u \neq v$. *Pythagoras's Theorem*, which seems to have been accidentally omitted from [VST], states that if u is orthogonal to v then $\|u+v\|^2 = \|u\|^2 + \|v\|^2$. The proof is straightforward:

$$\|u+v\|^2 = \langle (u+v), (u+v) \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \|u\|^2 + \|v\|^2$$

since orthogonality gives $\langle u, v \rangle = \langle v, u \rangle = 0$. It follows that if $u, v, w \in V$ and $u - v$ is orthogonal to $v - w$ then $d(u,w)^2 = d(u,v)^2 + d(v,w)^2$.

Now suppose that if P is the orthogonal projection of V onto a subspace U, and let $v \in V$. Then $P(v) \in U$ and $v - P(v)$ is orthogonal to all elements of U. If $x \in U$ is arbitrary then $P(v) - x \in U$, and so, by Pythagoras,

$$d(v,x)^2 = d(v, P(v))^2 + d(P(v), x)^2.$$

It follows that $d(v,x) \geq d(v, P(v))$, with equality if and only if $d(P(v), x) = 0$. But $d(P(v), x) = 0$ if and only if $x = P(v)$. Consequently $P(v)$ can be characterized as the element $x \in U$ for which $d(v,x)$ is minimal.