Summary of week 5 (lectures 13, 14 and 15)

Lectures 13 and 14 saw the conclusion of Chapter 4 of [VST]. First, Theorem 3.15 was proved, then Theorem 4.18. (The proofs can be found in [VST].) It follows from Theorem 4.18 that if v_1, v_2, …, v_n is a basis for the vector space V then

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \leftrightarrow \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

gives a one-to-one correspondence between F^n and V that preserves addition and scalar multiplication, in the following sense: if $\underset{\sim}{\lambda}$, $\underset{\sim}{\mu}$ are n-tuples (elements of F^n) and v, u the corresponding elements of V, then the element of V corresponding to the n-tuple $\underset{\sim}{\lambda} + \underset{\sim}{\mu}$ is $v + u$; similarly, if $\underset{\sim}{\lambda} \in F^n$ corresponds to $v \in V$ then for any scalar α the element of V corresponding to $\alpha\underset{\sim}{\lambda} \in F^n$ is $\alpha v \in V$. Thus V is isomorphic to F^n. (See the Week 4 summary for the definition of isomorphism and further comments. Note that the theorem stated at the end of the Week 4 summary is an immediate consequence of Theorem 4.18.)

The above process whereby a one-to-one correspondence is established between elements of V and n-tuples is totally analogous to the process by which the Euclidean plane is identified with $\mathbb{R}^2$ and three-dimensional Euclidean space is identified with $\mathbb{R}^3$. Choosing a basis coordinatizes V: each $v \in V$ is associated with an n-tuple of scalars, the coordinates of v, enabling us to identify V with F^n if we wish.

Definition. Let $\boldsymbol{b} = (v_1, v_2, \ldots . v_n)$ be a sequence of vectors that comprise a basis of the vector space V. For each $v \in V$ the *coordinate vector* of v relative to $\boldsymbol{b}$ is the n-tuple

$$\mathrm{cv}_{\boldsymbol{b}}(v) = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in F^n$$

such that $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$.

Note that a vector space will generally have more than one basis—the only exceptions to this are zero-dimensional vector spaces over any field and one-dimensional vector spaces over the two-element field—and consequently can be coordinatized in more than one way. We shall investigate relationships between different coordinate systems for the same space in due course. For the time being, the most important thing to be aware of is that once a basis is chosen, a finitely-generated vector space V can be viewed as a copy of F^n; the map $V \to F^n$ given by

$$v \mapsto \mathrm{cv}_{\boldsymbol{b}}(v) \qquad \text{(for all } v \in V)$$

and its inverse $F^n \to V$ given by

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

are isomorphisms.

Having completely classified finitely generated vector spaces, the next step is to classify linear transformations between finitely generated vector spaces. Since every finitely generated vector space over F is a copy of some F^n, it is enough to classify linear maps $F^n \to F^m$, where m, n are nonnegative integers. We have already noted that if A is any $m \times n$ matrix over F then multiplication by A determines a linear map from F^n to F^m. The classification theorem tells us that in fact these are the only linear maps from F^n to F^m.

Proposition. *Every linear transformation* $T\colon F^n \to F^m$ *is given by a formula of the form* $T(v) = Av$ *(for all* $v \in F^n$*) for some* $A \in \mathrm{Mat}(m \times n, F)$*.*

Proof. Let $T\colon F^n \to F^m$ be linear. Define scalars α_{ij}, for all $j \in \{1, 2, \ldots, n\}$ and $i \in \{1, 2, \ldots, m\}$, by the following formulas:

$$T\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{pmatrix}, \quad T\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{pmatrix}, \quad \ldots, \quad T\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{pmatrix}.$$

In other words, if e_j denotes the j-th vector in the standard basis of F^n, that is, the n-component column with 1 as its j-th entry and zeros elsewhere, then $T(e_j)$ is the m-component column whose i-th entry is α_{ij} for each i. By linearity of T we now find that for all values of the scalars λ_j,

$$\begin{aligned}
T\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} &= T\left(\lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \lambda_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right) \\
&= \lambda_1 T\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 T\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \lambda_n T\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{pmatrix} \lambda_1 + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{pmatrix} \lambda_2 + \cdots + \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{pmatrix} \lambda_n
\end{aligned}$$

$$= \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

That is, $T(v) = Av$ for all $v \in F^n$, where A is the $m \times n$ matrix whose (i, j)-entry is α_{ij}. $\square$

Definition. If $T: F^n \to F^m$ is a linear map then the *matrix of T relative to the standard bases of F^n and F^m* is the $m \times n$ matrix A such that $T(v) = Av$ for all $v \in F^n$.

Observe that the j-th column of the matrix of T is found by applying T to the j-th vector of the standard basis of F^n.

Corollary. *Let V be an n-dimensional vector space over F and W an m-dimensional vector space over F, and let $T: V \to W$ be a linear transformation. If $\boldsymbol{b}$ is any basis for V and $\boldsymbol{c}$ any basis for W then there is a unique $m \times n$ matrix A such that*

$$A\,\mathrm{cv}_{\boldsymbol{b}}(v) = \mathrm{cv}_{\boldsymbol{c}}(T(v)) \tag{1}$$

for all $v \in V$.

The idea here is that since V is a copy of F^n and W a copy of F^m, a linear map from V to W is essentially the same as a linear map from F^n to F^m, and as such corresponds to an $m \times n$ matrix A. The following diagram descibes the situation.

$$\begin{tikzcd}
V \arrow[d, phantom, "\cong" rotate=-90] & v \arrow[rr, mapsto, "T"] \arrow[d, leftrightarrow] & & T(v) \arrow[d, leftrightarrow] & W \arrow[d, phantom, "\cong" rotate=-90] \\
F^n & \mathrm{cv}_{\boldsymbol{b}}(v) \arrow[rr, mapsto, dashed] & & \mathrm{cv}_{\boldsymbol{c}}(T(v)) & F^m
\end{tikzcd}$$

An arbitrary $v \in V$ corresponds to an n-tuple $\mathrm{cv}_{\boldsymbol{b}}(v) \in F^n$, and T maps v to an element $T(v) \in W$, corresponding to the m-tuple $\mathrm{cv}_{\boldsymbol{c}}(T(v))$. Clearly the combined effect of $\mathrm{cv}_{\boldsymbol{b}}(v) \leftrightarrow v \mapsto T(v) \leftrightarrow \mathrm{cv}_{\boldsymbol{c}}(T(v))$ is a linear map from F^n toF^m, corresponding to the dashed arrow in the diagram. By the proposition proved above, this map is given by multiplication by some matrix A.

It is not hard to find how to compute the matrix A, by combining the isomorphisms $V \cong F^n$ and $W \cong F^m$ with the description of the matrix of a linear transformation $F^n \to F^m$ given above. Indeed, suppose that the basis $\boldsymbol{b}$ of V consists of the vectors $v_1, v_2, \dots, v_n$ and the basis $\boldsymbol{c}$ of W consists of the vectors $w_1, w_2, \dots, w_m$. The coordinate vectors (relative to $\boldsymbol{b}$) of $v_1, v_2, \dots, v_n$ are precisely the standard basis vectors of F^n; so the j-th column of A should be obtained by applying T to the j-th of these basis vectors, and finding the corresponding element of F^m. That is, the j-th column of A should be $\mathrm{cv}_{\boldsymbol{c}}(T(v_j))$.

It can be seen that the rule we have just given for computing A is consistent with Eq. (1) above. Indeed, putting $v = v_j$ and making use of the fact that $\mathrm{cv}_{\boldsymbol{b}}(v) = e_j$, Eq. (1) gives

$$Ae_j = \mathrm{cv}_{\boldsymbol{c}}(T(v)).$$

It is readily checked that Ae_j equals the j-th column of A: since the k-th entry of e_j is δ_{kj}, it follows that the i-th entry of Ae_j is $\sum_{k=1}^{n} A_{ik}\delta_{kj} = A_{ij}$, which is the i-th entry of the j-th column of A.

The following result is Theorem 2.9 of [VST]. The proof given in [VST] is based on the theory of row operations and elementary matrices; for variety, we give a proof based on the results of Chapter 4.

Proposition. *Let A, $B \in \mathrm{Mat}(n \times n, F)$. If $AB = I$ then $BA = I$.*

Proof. Observe first that $\mathrm{Mat}(n\times n, F)$ is a finitely generated vector space over F. Since an $n \times n$ matrix is essentially just an n^2-tuple of scalars, we can regard $\mathrm{Mat}(n \times n, F)$ as a copy of F^{n^2}, writing the elements as square arrays of scalars instead of columns of scalars. Identifying $\mathrm{Mat}(n \times n, F)$ and F^{n^2} in this way is compatible with the usual definitions of addition and scalar multiplication on both spaces. Thus $\mathrm{Mat}(n \times n, F)$ has dimension $d = n^2$ as a vector space over F; indeed, if for each pair of integers r, $s \in \{1, 2, \dots, n\}$ we define U_{rs} to be the matrix with (r, s)-entry equal to 1 and all other entries 0, then the n^2 matrices U_{rs} form a basis of $\mathrm{Mat}(n \times n, F)$.

For the purposes of the present proof, however, all that matters is that the dimension d of $\mathrm{Mat}(n \times n, F)$ is some finite number. The fact that $d = n^2$ is irrelevant.

Let A, $B \in \mathrm{Mat}(n\times n, F)$ be matrices satisfying $AB = I$. Consider the $d+1$ matrices I, A, A^2, …, A^d. Since $d + 1$ vectors in a d-dimensional vector space cannot be linearly independent (by Proposition 4.12 (ii)), there must exist scalars α_0, α_1, …, α_d which are not all zero and which satisfy

$$\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_d A^d = 0. \tag{2}$$

Choose $k \in \{0, 1, 2, \dots, d\}$ such that $\alpha_k \neq 0$; furthermore, choose k to be minimal with repect to this property, so that $\alpha_i = 0$ for all i such that $0 \leq i < k$. (Note that k could be zero, in which case there are no such i.) We can now write Eq. (2) as

$$\alpha_k A^k + \alpha_{k+1} A^{k+1} + \cdots + \alpha_d A^d = 0,$$

and using the fact that $\alpha_k \neq 0$ we can rearrange this to give

$$\begin{aligned} A^k &= -\alpha_k^{-1}(\alpha_{k+1}A^{k+1} + \alpha_{k+2}A^{k+2} + \cdots + \alpha_d A^d) \\ &= (\gamma_{k+1} I + \gamma_{k+2} A + \cdots + \gamma_d A^{d-k-1})A^{k+1} \end{aligned} \tag{3}$$

where $\gamma_i = -\alpha_k^{-1}\alpha_i$ (for each i).

The next step is to use induction on m show that $A^m B^m = I$ for all $m \geq 1$. The case $m = 1$ is trivial since we are given that $AB = I$. Suppose now that $m > 1$. The inductive hypothesis is that $A^{m-1}B^{m-1} = I$, and so

$$A^m B^m = AA^{m-1}B^{m-1}B = AIB = AB = I,$$

as required. Now multiplying Eq. (3) on the right by B^k gives

$$\begin{aligned} I &= A^k B^k \\ &= (\gamma_{k+1}I + \gamma_{k+2}A + \cdots + \gamma_d A^{d-k-1})A^{k+1}B^k \\ &= (\gamma_{k+1}I + \gamma_{k+2}A + \cdots + \gamma_d A^{d-k-1})AA^k B^k \\ &= (\gamma_{k+1}I + \gamma_{k+2}A + \cdots + \gamma_d A^{d-k-1})A, \end{aligned}$$

and multiplying on the right by B again gives

$$\begin{aligned} B &= (\gamma_{k+1}I + \gamma_{k+2}A + \cdots + \gamma_d A^{d-k-1})AB \\ &= \gamma_{k+1}I + \gamma_{k+2}A + \cdots + \gamma_d A^{d-k-1}. \end{aligned}$$

But this gives

$$\begin{aligned} BA &= (\gamma_{k+1}I + \gamma_{k+2}A + \cdots + \gamma_d A^{d-k-1})A \\ &= \gamma_{k+1}A + \gamma_{k+2}A^2 + \cdots + \gamma_d A^{d-k} \\ &= A(\gamma_{k+1}I + \gamma_{k+2}A + \cdots + \gamma_d A^{d-k-1}) \\ &= AB. \end{aligned}$$

Since $AB = I$ it follows that $BA = I$, as required. □

Example. Define v_1, v_2, $v_3 \in \mathbb{R}^3$ by

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

We shall show that the vectors v_1, v_2, v_3 comprise a basis $\boldsymbol{b}$ of $\mathbb{R}^3$, and calculate $\mathrm{cv}_{\boldsymbol{b}}(e_1)$, $\mathrm{cv}_{\boldsymbol{b}}(e_2)$ and $\mathrm{cv}_{\boldsymbol{b}}(e_3)$,the coordinate vectors relative to $\boldsymbol{b}$ of the three vectors that comprise the standard basis of $\mathbb{R}^3$:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Note that this example is similar to #2 on p. 96 of [VST].

Since $\mathbb{R}^3$ has dimension 3 it follows 4.12 (iii) that v_1, v_2, v_3 are linearly independent if and only if they span $\mathbb{R}^3$. Now the equation

$$\begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix} = \lambda_1\begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda_2\begin{pmatrix}0\\1\\0\end{pmatrix} + \lambda_3\begin{pmatrix}0\\0\\1\end{pmatrix}$$

shows that every element of $\mathbb{R}^3$ can be expressed as a linear combination of e_1, e_2, e_3, and if we can express each of e_1, e_2, e_3 as linear combinations of v_1, v_2, v_3 then it will follow that every element of $\mathbb{R}^3$ is expressible as a linear combination of v_1, v_2, v_3. Thus it will follow that v_1, v_2, v_3 span $\mathbb{R}^3$, and hence are linearly independent also. So if we can solve the equations

$$\left.\begin{aligned}
\alpha_1\begin{pmatrix}1\\1\\1\end{pmatrix} + \beta_1\begin{pmatrix}1\\0\\1\end{pmatrix} + \gamma_1\begin{pmatrix}1\\2\\2\end{pmatrix} &= \begin{pmatrix}1\\0\\0\end{pmatrix}\\
\alpha_2\begin{pmatrix}1\\1\\1\end{pmatrix} + \beta_2\begin{pmatrix}1\\0\\1\end{pmatrix} + \gamma_2\begin{pmatrix}1\\2\\2\end{pmatrix} &= \begin{pmatrix}0\\1\\0\end{pmatrix}\\
\alpha_3\begin{pmatrix}1\\1\\1\end{pmatrix} + \beta_3\begin{pmatrix}1\\0\\1\end{pmatrix} + \gamma_3\begin{pmatrix}1\\2\\2\end{pmatrix} &= \begin{pmatrix}0\\0\\1\end{pmatrix}
\end{aligned}\right\}\qquad (4)$$

then it will follow that v_1, v_2, v_3 is a basis; moreover, it will also follow that

$$\mathrm{cv}_{\boldsymbol{b}}(e_1) = \begin{pmatrix}\alpha_1\\ \beta_1\\ \gamma_1\end{pmatrix}, \qquad \mathrm{cv}_{\boldsymbol{b}}(e_2) = \begin{pmatrix}\alpha_2\\ \beta_2\\ \gamma_2\end{pmatrix}, \qquad \mathrm{cv}_{\boldsymbol{b}}(e_3) = \begin{pmatrix}\alpha_3\\ \beta_3\\ \gamma_3\end{pmatrix}.$$

Now the three systems (4) can be combined into the single matrix equation

$$\begin{pmatrix}1&1&1\\1&0&2\\1&1&2\end{pmatrix}\begin{pmatrix}\alpha_1&\alpha_2&\alpha_3\\ \beta_1&\beta_2&\beta_3\\ \gamma_1&\gamma_2&\gamma_3\end{pmatrix} = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

and so the desired coordinate vectors are the columns of the inverse of the matrix

$$A = \begin{pmatrix}1&1&1\\1&0&2\\1&1&2\end{pmatrix}.$$

(Recall that, by the proposition proved above, a matrix B satisfying $AB = I$ will also satisfy $BA = I$, and hence be the inverse of A.)

The inverse of A is best found by the row operations method you were taught in 1st year.

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \xrightarrow[R_3:=R_3-R_1]{R_2:=R_2-R_1} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right)$$
$$\xrightarrow[R_2:=R_2-R_3]{R_1:=R_1-R_3} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & 0 & -1 \\ 0 & -1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right)$$
$$\xrightarrow{R_1:=R_1+R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 1 & -2 \\ 0 & -1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right)$$
$$\xrightarrow{R_2:=-R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 1 & -2 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right).$$

Therefore $\mathrm{cv}_{\boldsymbol{b}}(e_1) = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$, $\mathrm{cv}_{\boldsymbol{b}}(e_2) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ and $\mathrm{cv}_{\boldsymbol{b}}(e_3) = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$.

We can apply the same reasoning to any sequence of n vectors $v_1, v_2, \ldots, v_n$ in F^n, for any positive integer n and any field F. The conclusion is that if we let $A = (v_1|v_2|\ldots|v_n)$, the $n \times n$ matrix whose columns are the given vectors $v_1, v_2, \ldots, v_n$, then the following conditions are equivalent:

(1) $\mathrm{CS}(A) = F^n$; that is, $v_1, v_2, \ldots, v_n$ span F^n.
(2) For every $v \in F^n$ the equations $Ax = v$ have a solution $x \in F^n$.
(3) A $n \times n$ matrix X exists such that $AX = I$.
(4) A has an inverse.
(5) $v_1, v_2, \ldots, v_n$ are linearly independent.
(6) $\mathrm{RN}(A) = \{0\}$; that is, the only solution of $Ax = 0$ (with $x \in F^n$) is $x = 0$.
(7) A $n \times n$ matrix X exists such that $XA = I$.
(8) $\mathrm{RS}(A) = {}^{\mathrm{t}}(F^n)$ (the space of all n-component row vectors).
(9) For every $v \in {}^{\mathrm{t}}(F^n)$ the equations $xA = v$ have a solution $x \in {}^{\mathrm{t}}(F^n)$.
(10) $\mathrm{LN}(A) = \{0\}$; that is, the only solution of $xA = 0$ (with $x \in {}^{\mathrm{t}}(F^n)$) is $x = 0$.

We shall add to this list later: see 8.20.1 of [VST].

In Lecture 15 we commenced Chapter 5 of [VST], dealing with inner product spaces. In this section of the course it is necessary to assume that the field of scalars is either $\mathbb{R}$ (the real numbers) or $\mathbb{C}$ (the complex numbers).

We have seen how the Euclidean plane and three-dimensional Euclidean space can be regarded as vector spaces over $\mathbb{R}$; moreover, Cartesian coordinates permit us to identify the Euclidean plane with $\mathbb{R}^2$ and three-dimensional Euclidean space with $\mathbb{R}^3$. However, the general theory of vector spaces ignores the two most fundamental concepts of Euclidean geometry: the *length* of a line segment and the *angle* between two lines. It is natural to attempt to also use coordinates to analyse these concepts.

If we choose perpendicular coordinate axes, in the customary manner, then the length of the line segment OP joining the origin O to the point P with coordinates $\binom{x_1}{x_2}$ is given by $|OP| = \sqrt{x_1^2 + x_2^2}$. If Q has coordinates $\binom{y_1}{y_2}$ then (similarly) $|OQ| = \sqrt{y_1^2 + y_2^2}$, and $|PQ| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$. Now an application of the cosine rule in the triangle OPQ gives

$$|OP|\,|OQ|\,\cos\theta = \tfrac{1}{2}(|OP|^2 + |OQ|^2 - |PQ|^2) = x_1y_1 + x_2y_2,$$

where θ is the angle $\angle POQ$. A similar calculation works in $\mathbb{R}^3$; see [VST] p. 99.

Proceeding by analogy, in $\mathbb{R}^n$ let us define the *length* (also called the *norm*) of a vector

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

to be the scalar $\|v\|$ given by

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

If P and Q are points whose coordinates are ${}^{\mathrm{t}}(x_1, \ldots, x_n)$ and ${}^{\mathrm{t}}(y_1, \ldots, y_n)$ then there will be a plane through the origin (a two-dimensional subspace of $\mathbb{R}^n$) containing both these points, and applying the cosine rule in the triangle OPQ as above gives

$$|OP|\,|OQ|\,\cos\theta = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Definition. The *dot product* on $\mathbb{R}^n$ is the function $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Observe that the length of a vector $v \in \mathbb{R}^n$ is given by $\|v\| = \sqrt{v \cdot v}$; hence, by the formulas derived above, the angle θ between vectors u, $v \in \mathbb{R}^n$ is given by $\theta = \arccos\bigl((u \cdot v)/\sqrt{(u \cdot u)(v \cdot v)}\bigr)$. So the dot product enables us to deal with both lengths and angles in $\mathbb{R}^n$.

In the lecture the proofs were sketched of the three main basic properties of the dot product (see [VST] p. 100), and the definition of a *real inner product space* was given (see [VST] 5.1). It was also explained how one may define an inner product on the space $\mathscr{C}[a, b]$ of all continuous real-valued functions on the closed interval $[a, b]$ (see [VST] p. 101). Finally, the definition of the complex dot product on complex n-space $\mathbb{C}^n$ was given (see [VST] p.104).