Chapter 8

The Dual Space, Duality

8.1 The Dual Space $E^*$ and Linear Forms

In Section 1.7 we defined linear forms, the dual space $E^* = \text{Hom}(E, K)$ of a vector space $E$, and showed the existence of dual bases for vector spaces of finite dimension.

In this chapter, we take a deeper look at the connection between a space $E$ and its dual space $E^*$.

As we will see shortly, every linear map $f : E \rightarrow F$ gives rise to a linear map $f^\top : F^* \rightarrow E^*$, and it turns out that in a suitable basis, the matrix of $f^\top$ is the transpose of the matrix of $f$. 
Thus, the notion of dual space provides a conceptual explanation of the phenomena associated with transposition.

But it does more, because it allows us to view subspaces as solutions of sets of linear equations and vice-versa.

Consider the following set of two “linear equations” in $\mathbb{R}^3$,

\[
\begin{align*}
x - y + z &= 0 \\
x - y - z &= 0,
\end{align*}
\]

and let us find out what is their set $V$ of common solutions $(x, y, z) \in \mathbb{R}^3$.

By subtracting the second equation from the first, we get $2z = 0$, and by adding the two equations, we find that $2(x - y) = 0$, so the set $V$ of solutions is given by

\[
\begin{align*}
y &= x \\
z &= 0.
\end{align*}
\]
This is a one dimensional subspace of $\mathbb{R}^3$. Geometrically, this is the line of equation $y = x$ in the plane $z = 0$.

Now, why did we say that the above equations are linear?

This is because, as functions of $(x, y, z)$, both maps $f_1: (x, y, z) \mapsto x - y + z$ and $f_2: (x, y, z) \mapsto x - y - z$ are linear.

The set of all such linear functions from $\mathbb{R}^3$ to $\mathbb{R}$ is a vector space; we used this fact to form linear combinations of the “equations” $f_1$ and $f_2$.

Observe that the dimension of the subspace $V$ is 1.

The ambient space has dimension $n = 3$ and there are two “independent” equations $f_1, f_2$, so it appears that the dimension $\dim(V)$ of the subspace $V$ defined by $m$ independent equations is

$$\dim(V) = n - m,$$

which is indeed a general fact (proved in Theorem 8.1).
More generally, in $\mathbb{R}^n$, a linear equation is determined by an $n$-tuple $(a_1, \ldots, a_n) \in \mathbb{R}^n$, and the solutions of this linear equation are given by the $n$-tuples $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that

$$a_1x_1 + \cdots + a_nx_n = 0;$$

these solutions constitute the kernel of the linear map

$$(x_1, \ldots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n.$$

The above considerations assume that we are working in the canonical basis $(e_1, \ldots, e_n)$ of $\mathbb{R}^n$, but we can define “linear equations” independently of bases and in any dimension, by viewing them as elements of the vector space $\text{Hom}(E, K)$ of linear maps from $E$ to the field $K$. 
Definition 8.1. Given a vector space $E$, the vector space $\text{Hom}(E, K)$ of linear maps from $E$ to $K$ is called the \textit{dual space (or dual)} of $E$. The space $\text{Hom}(E, K)$ is also denoted by $E^*$, and the linear maps in $E^*$ are called \textit{the linear forms}, or \textit{covectors}. The dual space $E^{**}$ of the space $E^*$ is called the \textit{bidual} of $E$.

As a matter of notation, linear forms $f : E \to K$ will also be denoted by starred symbol, such as $u^*$, $x^*$, etc.
8.2 Pairing and Duality Between $E$ and $E^*$

Given a linear form $u^* \in E^*$ and a vector $v \in E$, the result $u^*(v)$ of applying $u^*$ to $v$ is also denoted by $\langle u^*, v \rangle$.

This defines a binary operation $\langle - , - \rangle : E^* \times E \rightarrow K$ satisfying the following properties:

$$\langle u_1^* + u_2^*, v \rangle = \langle u_1^*, v \rangle + \langle u_2^*, v \rangle$$
$$\langle u^*, v_1 + v_2 \rangle = \langle u^*, v_1 \rangle + \langle u^*, v_2 \rangle$$
$$\langle \lambda u^*, v \rangle = \lambda \langle u^*, v \rangle$$
$$\langle u^*, \lambda v \rangle = \lambda \langle u^*, v \rangle.$$

The above identities mean that $\langle - , - \rangle$ is a \textit{bilinear map}, since it is linear in each argument.

It is often called the \textit{canonical pairing} between $E^*$ and $E$. 
In view of the above identities, given any fixed vector $v \in E$, the map $\text{eval}_v : E^* \to K$ (evaluation at $v$) defined such that

$$\text{eval}_v(u^*) = \langle u^*, v \rangle = u^*(v) \quad \text{for every } u^* \in E^*$$

is a linear map from $E^*$ to $K$, that is, $\text{eval}_v$ is a linear form in $E^{**}$.

Again from the above identities, the map $\text{eval}_E : E \to E^{**}$, defined such that

$$\text{eval}_E(v) = \text{eval}_v \quad \text{for every } v \in E,$$

is a linear map.

We shall see that it is injective, and that it is an isomorphism when $E$ has finite dimension.
We now formalize the notion of the set $V^0$ of linear equations vanishing on all vectors in a given subspace $V \subseteq E$, and the notion of the set $U^0$ of common solutions of a given set $U \subseteq E^*$ of linear equations.

The duality theorem (Theorem 8.1) shows that the dimensions of $V$ and $V^0$, and the dimensions of $U$ and $U^0$, are related in a crucial way.

It also shows that, in finite dimension, the maps $V \mapsto V^0$ and $U \mapsto U^0$ are inverse bijections from subspaces of $E$ to subspaces of $E^*$. 
Definition 8.2. Given a vector space $E$ and its dual $E^*$, we say that a vector $v \in E$ and a linear form $u^* \in E^*$ are \textit{orthogonal} iff $\langle u^*, v \rangle = 0$. Given a subspace $V$ of $E$ and a subspace $U$ of $E^*$, we say that $V$ and $U$ are orthogonal iff $\langle u^*, v \rangle = 0$ for every $u^* \in U$ and every $v \in V$. Given a subset $V$ of $E$ (resp. a subset $U$ of $E^*$), the \textit{orthogonal} $V^0$ of $V$ is the subspace $V^0$ of $E^*$ defined such that

$$V^0 = \{u^* \in E^* \mid \langle u^*, v \rangle = 0, \text{ for every } v \in V\}$$

(resp. the \textit{orthogonal} $U^0$ of $U$ is the subspace $U^0$ of $E$ defined such that

$$U^0 = \{v \in E \mid \langle u^*, v \rangle = 0, \text{ for every } u^* \in U\}$$

The subspace $V^0 \subseteq E^*$ is also called the \textit{annihilator} of $V$.\)
The subspace $U^0 \subseteq E$ annihilated by $U \subseteq E^*$ does not have a special name. It seems reasonable to call it the \textit{linear subspace (or linear variety) defined by $U$}.

Informally, $V^0$ is the \textit{set of linear equations that vanish on $V$}, and $U^0$ is the \textit{set of common zeros of all linear equations in $U$}. We can also define $V^0$ by

$$V^0 = \{ u^* \in E^* \mid V \subseteq \text{Ker } u^* \}$$

and $U^0$ by

$$U^0 = \bigcap_{u^* \in U} \text{Ker } u^*.$$ 

Observe that $E^0 = \{0\} = (0)$, and $\{0\}^0 = E^*$. Furthermore, if $V_1 \subseteq V_2 \subseteq E$, then $V_2^0 \subseteq V_1^0 \subseteq E^*$, and if $U_1 \subseteq U_2 \subseteq E^*$, then $U_2^0 \subseteq U_1^0 \subseteq E$. 

It can also be shown that that $V \subseteq V^{00}$ for every subspace $V$ of $E$, and that $U \subseteq U^{00}$ for every subspace $U$ of $E^*$.

We will see shortly that in finite dimension, we have

$$V = V^{00} \quad \text{and} \quad U = U^{00}.$$  

Here are some examples. Let $E = M_2(\mathbb{R})$, the space of real $2 \times 2$ matrices, and let $V$ be the subspace of $M_2(\mathbb{R})$ spanned by the matrices

$$
\begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}, \quad 
\begin{pmatrix}
1 & 0 \\
0 & 0
\end{pmatrix}, \quad 
\begin{pmatrix}
0 & 0 \\
0 & 1
\end{pmatrix}.
$$

We check immediately that the subspace $V$ consists of all matrices of the form

$$
\begin{pmatrix}
b & a \\
a & c
\end{pmatrix},
$$

that is, all symmetric matrices.
The matrices
\[
\begin{pmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{pmatrix}
\]
in \(V\) satisfy the equation
\[a_{12} - a_{21} = 0,\]
and all scalar multiples of these equations, so \(V^0\) is the subspace of \(E^*\) spanned by the linear form given by
\[u^*(a_{11}, a_{12}, a_{21}, a_{22}) = a_{12} - a_{21}.\]

By the duality theorem (Theorem 8.1) we have
\[
\dim(V^0) = \dim(E) - \dim(V) = 4 - 3 = 1.
\]

The above example generalizes to \(E = M_n(\mathbb{R})\) for any \(n \geq 1\), but this time, consider the space \(U\) of linear forms asserting that a matrix \(A\) is symmetric; these are the linear forms spanned by the \(n(n - 1)/2\) equations
\[a_{ij} - a_{ji} = 0, \quad 1 \leq i < j \leq n;\]
Note there are no constraints on diagonal entries, and half of the equations

\[ a_{ij} - a_{ji} = 0, \quad 1 \leq i \neq j \leq n \]

are redundant. It is easy to check that the equations (linear forms) for which \( i < j \) are linearly independent.

To be more precise, let \( U \) be the space of linear forms in \( E^* \) spanned by the linear forms

\[
\begin{align*}
  u_{ij}^*(a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{n1}, \ldots, a_{nn}) &= a_{ij} - a_{ji}, \quad 1 \leq i < j \leq n.
\end{align*}
\]

The dimension of \( U \) is \( n(n - 1)/2 \). Then, the set \( U^0 \) of common solutions of these equations is the space \( S(n) \) of symmetric matrices.

By the duality theorem (Theorem 8.1), this space has dimension

\[
\frac{n(n + 1)}{2} = n^2 - \frac{n(n - 1)}{2}.
\]
If $E = M_n(\mathbb{R})$, consider the subspace $U$ of linear forms in $E^*$ spanned by the linear forms

$$u_{ij}^*(a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{n1}, \ldots, a_{nn}) = a_{ij} + a_{ji}, \quad 1 \leq i < j \leq n$$

$$u_{ii}^*(a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{n1}, \ldots, a_{nn}) = a_{ii}, \quad 1 \leq i \leq n.$$

It is easy to see that these linear forms are linearly independent, so $\dim(U) = n(n + 1)/2$.

The space $U^0$ of matrices $A \in M_n(\mathbb{R})$ satisfying all of the above equations is clearly the space $\text{Skew}(n)$ of skew-symmetric matrices.

By the duality theorem (Theorem 8.1), the dimension of $U^0$ is

$$\frac{n(n - 1)}{2} = n^2 - \frac{n(n + 1)}{2}.$$
For yet another example with $E = M_n(\mathbb{R})$, for any $A \in M_n(\mathbb{R})$, consider the linear form in $E^*$ given by

$$\text{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn},$$

called the *trace* of $A$.

The subspace $U^0$ of $E$ consisting of all matrices $A$ such that $\text{tr}(A) = 0$ is a space of dimension $n^2 - 1$.

The dimension equations

$$\dim(V) + \dim(V^0) = \dim(E)$$

$$\dim(U) + \dim(U^0) = \dim(E)$$

are always true (if $E$ is finite-dimensional). This is part of the duality theorem (Theorem 8.1).

In contrast with the previous examples, given a matrix $A \in M_n(\mathbb{R})$, the equations asserting that $A^\top A = I$ are not linear constraints.
For example, for $n = 2$, we have
\[
\begin{align*}
a_{11}^2 + a_{21}^2 &= 1 \\
a_{21}^2 + a_{22}^2 &= 1 \\
a_{11}a_{12} + a_{21}a_{22} &= 0.
\end{align*}
\]

Given a vector space $E$ and any basis $(u_i)_{i \in I}$ for $E$, we can associate to each $u_i$ a linear form $u_i^* \in E^*$, and the $u_i^*$ have some remarkable properties.

**Definition 8.3.** Given a vector space $E$ and any basis $(u_i)_{i \in I}$ for $E$, by Proposition 1.11, for every $i \in I$, there is a unique linear form $u_i^*$ such that
\[
u_i^*(u_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}
\]
for every $j \in I$. The linear form $u_i^*$ is called the *coordinate form* of index $i$ w.r.t. the basis $(u_i)_{i \in I}$. 
The reason for the terminology *coordinate form* was explained in Section 1.7.

We proved in Theorem 1.14 that if \((u_1, \ldots, u_n)\) is a basis of \(E\), then \((u_1^*, \ldots, u_n^*)\) is a basis of \(E^*\) called the *dual basis*.

If \((u_1, \ldots, u_n)\) is a basis of \(\mathbb{R}^n\) (more generally \(K^n\)), it is possible to find explicitly the dual basis \((u_1^*, \ldots, u_n^*)\), where each \(u_i^*\) is represented by a row vector.
For example, consider the columns of the Bézier matrix

\[
B_4 = \begin{pmatrix}
1 & -3 & 3 & -1 \\
0 & 3 & -6 & 3 \\
0 & 0 & 3 & -3 \\
0 & 0 & 0 & 1
\end{pmatrix}.
\]

The form \( u_1^* \) is represented by a row vector \( (\lambda_1 \, \lambda_2 \, \lambda_3 \, \lambda_4) \) such that

\[
(\lambda_1 \, \lambda_2 \, \lambda_3 \, \lambda_4) \begin{pmatrix}
1 & -3 & 3 & -1 \\
0 & 3 & -6 & 3 \\
0 & 0 & 3 & -3 \\
0 & 0 & 0 & 1
\end{pmatrix} = (1 \, 0 \, 0 \, 0).
\]

This implies that \( u_1^* \) is the first row of the inverse of \( B_4 \).
Since

\[ B_4^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1/3 & 2/3 & 1 \\ 0 & 0 & 1/3 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \]

the linear forms \((u_1^*, u_2^*, u_3^*, u_4^*)\) correspond to the rows of \(B_4^{-1}\).

In particular, \(u_1^*\) is represented by \((1 1 1 1)\).

The above method works for any \(n\). Given any basis \((u_1, \ldots, u_n)\) of \(\mathbb{R}^n\), if \(P\) is the \(n \times n\) matrix whose \(j\)th column is \(u_j\), then the dual form \(u_i^*\) is given by the \(i\)th row of the matrix \(P^{-1}\).

We have the following important duality theorem adapted from E. Artin.
8.3 The Duality Theorem

Theorem 8.1. (Duality theorem) Let $E$ be a vector space of dimension $n$. The following properties hold:

(a) For every basis $(u_1, \ldots, u_n)$ of $E$, the family of coordinate forms $(u_1^*, \ldots, u_n^*)$ is a basis of $E^*$.

(b) For every subspace $V$ of $E$, we have $V^{00} = V$.

(c) For every pair of subspaces $V$ and $W$ of $E$ such that $E = V \oplus W$, with $V$ of dimension $m$, for every basis $(u_1, \ldots, u_n)$ of $E$ such that $(u_1, \ldots, u_m)$ is a basis of $V$ and $(u_{m+1}, \ldots, u_n)$ is a basis of $W$, the family $(u_1^*, \ldots, u_m^*)$ is a basis of the orthogonal $W^0$ of $W$ in $E^*$. Furthermore, we have $W^{00} = W$, and

$$\dim(W) + \dim(W^0) = \dim(E).$$

(d) For every subspace $U$ of $E^*$, we have

$$\dim(U) + \dim(U^0) = \dim(E),$$

where $U^0$ is the orthogonal of $U$ in $E$, and $U^{00} = U$. 
8.3. THE DUALITY THEOREM

Part (a) of Theorem 8.1 shows that

\[ \dim(E) = \dim(E^*), \]

and if \((u_1, \ldots, u_n)\) is a basis of \(E\), then \((u_1^*, \ldots, u_n^*)\) is a basis of the dual space \(E^*\) called the dual basis of \((u_1, \ldots, u_n)\).

Define the function \(\mathcal{E}\) from subspaces of \(E\) to subspaces of \(E^*\) and the function \(\mathcal{Z}\) from subspaces of \(E^*\) to subspaces of \(E\) by

\[
\mathcal{E}(V) = V^0, \quad V \subseteq E \\
\mathcal{Z}(U) = U^0, \quad U \subseteq E^*.
\]

By part (c) and (d) of theorem 8.1,

\[
(\mathcal{Z} \circ \mathcal{E})(V) = V^{00} = V \\
(\mathcal{E} \circ \mathcal{Z})(U) = U^{00} = U,
\]

so \(\mathcal{Z} \circ \mathcal{E} = \text{id}\) and \(\mathcal{E} \circ \mathcal{Z} = \text{id}\), and the maps \(\mathcal{E}\) and \(\mathcal{V}\) are inverse bijections.
These maps set up a *duality* between subspaces of $\mathbf{E}$, and subspaces of $\mathbf{E}^*$. 

One should be careful that this bijection does not hold if $\mathbf{E}$ has infinite dimension. Some restrictions on the dimensions of $\mathbf{U}$ and $\mathbf{V}$ are needed.

Suppose that $\mathbf{V}$ is a subspace of $\mathbb{R}^n$ of dimension $m$ and that $(v_1, \ldots, v_m)$ is a basis of $\mathbf{V}$.

To find a basis of $V^0$, we first extend $(v_1, \ldots, v_m)$ to a basis $(v_1, \ldots, v_n)$ of $\mathbb{R}^n$, and then by part (c) of Theorem 8.1, we know that $(v_{m+1}^*, \ldots, v_n^*)$ is a basis of $V^0$.

For example, suppose that $\mathbf{V}$ is the subspace of $\mathbb{R}^4$ spanned by the two linearly independent vectors

$$
v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix},
$$

the first two vectors of the Haar basis in $\mathbb{R}^4$. 
8.3. THE DUALITY THEOREM

The four columns of the Haar matrix

\[ W = \begin{pmatrix}
1 & 1 & 1 & 0 \\
1 & 1 & -1 & 0 \\
1 & -1 & 0 & 1 \\
1 & -1 & 0 & -1 \\
\end{pmatrix} \]

form a basis of \( \mathbb{R}^4 \), and the inverse of \( W \) is given by

\[ W^{-1} = \begin{pmatrix}
1/4 & 1/4 & 1/4 & 1/4 \\
1/4 & 1/4 & -1/4 & -1/4 \\
1/2 & -1/2 & 0 & 0 \\
0 & 0 & 1/2 & -1/2 \\
\end{pmatrix}. \]

Since the dual basis \( (v_1^*, v_2^*, v_3^*, v_4^*) \) is given by the row of \( W^{-1} \), the last two rows of \( W^{-1} \),

\[ \begin{pmatrix}
1/2 & -1/2 & 0 & 0 \\
0 & 0 & 1/2 & -1/2 \\
\end{pmatrix}, \]

form a basis of \( V^0 \).
We also obtain a basis by rescaling by the factor $1/2$, so the linear forms given by the row vectors

\[
\begin{pmatrix}
1 & -1 & 0 & 0 \\
0 & 0 & 1 & -1
\end{pmatrix}
\]

form a basis of $V^0$, the space of linear forms (linear equations) that vanish on the subspace $V$.

The method that we described to find $V^0$ requires first extending a basis of $V$ and then inverting a matrix, but there is a more direct method.
Indeed, let \( A \) be the \( n \times m \) matrix whose columns are the basis vectors \( (v_1, \ldots, v_m) \) of \( V \). Then, a linear form \( u \) represented by a row vector belongs to \( V^0 \) iff \( uv_i = 0 \) for \( i = 1, \ldots, m \) iff

\[
u A = 0
\]

iff

\[
A^\top u^\top = 0.
\]

Therefore, all we need to do is to \textit{find a basis of the nullspace of} \( A^\top \).

This can be done quite effectively using the reduction of a matrix to reduced row echelon form (rref); see Section 5.9.
Here is another example illustrating the power of Theorem 8.1.

Let $E = \text{M}_n(\mathbb{R})$, and consider the equations asserting that the sum of the entries in every row of a matrix $\in \text{M}_n(\mathbb{R})$ is equal to the same number.

We have $n - 1$ equations

$$\sum_{j=1}^{n}(a_{ij} - a_{i+1j}) = 0, \quad 1 \leq i \leq n - 1,$$

and it is easy to see that they are linearly independent.

Therefore, the space $U$ of linear forms in $E^*$ spanned by the above linear forms (equations) has dimension $n - 1$, and the space $U^0$ of matrices satisfying all these equations has dimension $n^2 - n + 1$.

It is not so obvious to find a basis for this space.
When $E$ is of finite dimension $n$ and $(u_1, \ldots, u_n)$ is a basis of $E$, we noted that the family $(u_1^*, \ldots, u_n^*)$ is a basis of the dual space $E^*$,

Let us see how the coordinates of a linear form $\varphi^* \in E^*$ over the basis $(u_1^*, \ldots, u_n^*)$ vary under a change of basis.

Let $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ be two bases of $E$, and let $P = (a_{ij})$ be the change of basis matrix from $(u_1, \ldots, u_n)$ to $(v_1, \ldots, v_n)$, so that

$$v_j = \sum_{i=1}^{n} a_{ij} u_i.$$ 

If

$$\varphi^* = \sum_{i=1}^{n} \varphi_i u_i^* = \sum_{i=1}^{n} \varphi_i' v_i^*,$$

after some algebra, we get

$$\varphi_j' = \sum_{i=1}^{n} a_{ij} \varphi_i.$$
Comparing with the change of basis

\[ v_j = \sum_{i=1}^{n} a_{i,j} u_i, \]

we note that this time, the coordinates \((\varphi_i)\) of the linear form \(\varphi^*\) change in the \textit{same direction} as the change of basis.

For this reason, we say that the coordinates of linear forms are \textit{covariant}.

By abuse of language, it is often said that linear forms are \textit{covariant}, which explains why the term \textit{covector} is also used for a linear form.

Observe that if \((e_1, \ldots, e_n)\) is a basis of the vector space \(E\), then, as a linear map from \(E\) to \(K\), every linear form \(f \in E^*\) is represented by a \(1 \times n\) matrix, that is, by a \textit{row vector}

\[ (\lambda_1 \cdots \lambda_n), \]

with respect to the basis \((e_1, \ldots, e_n)\) of \(E\), and 1 of \(K\), where \(f(e_i) = \lambda_i\).
A vector \( u = \sum_{i=1}^{n} u_i e_i \in E \) is represented by a \( n \times 1 \) matrix, that is, by a \textit{column vector} 

\[
\begin{pmatrix}
  u_1 \\
  \vdots \\
  u_n
\end{pmatrix},
\]

and the action of \( f \) on \( u \), namely \( f(u) \), is represented by the matrix product

\[
(\lambda_1 \cdots \lambda_n) \begin{pmatrix}
  u_1 \\
  \vdots \\
  u_n
\end{pmatrix} = \lambda_1 u_1 + \cdots + \lambda_n u_n.
\]

On the other hand, with respect to the dual basis \((e^*_1, \ldots, e^*_n)\) of \( E^* \), the linear form \( f \) is represented by the column vector

\[
\begin{pmatrix}
  \lambda_1 \\
  \vdots \\
  \lambda_n
\end{pmatrix}.
\]
We will now pin down the relationship between a vector space $E$ and its bidual $E^{**}$.

**Proposition 8.2.** Let $E$ be a vector space. The following properties hold:

(a) The linear map $\text{eval}_E: E \to E^{**}$ defined such that

$$\text{eval}_E(v) = \text{eval}_v, \text{ for all } v \in E,$$

that is, $\text{eval}_E(v)(u^*) = \langle u^*, v \rangle = u^*(v)$ for every $u^* \in E^*$, is injective.

(b) When $E$ is of finite dimension $n$, the linear map $\text{eval}_E: E \to E^{**}$ is an isomorphism (called the canonical isomorphism).

When $E$ is of finite dimension and $(u_1, \ldots, u_n)$ is a basis of $E$, in view of the canonical isomorphism $\text{eval}_E: E \to E^{**}$, the basis $(u_1^{**}, \ldots, u_n^{**})$ of the bidual is identified with $(u_1, \ldots, u_n)$.

Proposition 8.2 can be reformulated very fruitfully in terms of pairings.
Definition 8.4. Given two vector spaces $E$ and $F$ over $K$, a **pairing between $E$ and $F$** is a bilinear map

$\varphi: E \times F \to K$. Such a pairing is **nondegenerate** iff

1. For every $u \in E$, if $\varphi(u, v) = 0$ for all $v \in F$, then $u = 0$, and

2. For every $v \in F$, if $\varphi(u, v) = 0$ for all $u \in E$, then $v = 0$.

A pairing $\varphi: E \times F \to K$ is often denoted by $\langle -, - \rangle: E \times F \to K$.

For example, the map $\langle -, - \rangle: E^* \times E \to K$ defined earlier is a nondegenerate pairing (use the proof of (a) in Proposition 8.2).

Given a pairing $\varphi: E \times F \to K$, we can define two maps

$l_\varphi: E \to F^*$ and $r_\varphi: F \to E^*$ as follows:
For every $u \in E$, we define the linear form $l_\varphi(u)$ in $F^*$ such that

$$l_\varphi(u)(y) = \varphi(u, y) \quad \text{for every } y \in F,$$

and for every $v \in F$, we define the linear form $r_\varphi(v)$ in $E^*$ such that

$$r_\varphi(v)(x) = \varphi(x, v) \quad \text{for every } x \in E.$$

**Proposition 8.3.** Given two vector spaces $E$ and $F$ over $K$, for every nondegenerate pairing

$\varphi: E \times F \to K$ between $E$ and $F$, the maps

$l_\varphi: E \to F^*$ and $r_\varphi: F \to E^*$ are linear and injective. Furthermore, if $E$ and $F$ have finite dimension, then this dimension is the same and $l_\varphi: E \to F^*$ and $r_\varphi: F \to E^*$ are bijections.

When $E$ has finite dimension, the nondegenerate pairing $\langle -, - \rangle: E^* \times E \to K$ yields another proof of the existence of a natural isomorphism between $E$ and $E^{**}$.

Interesting nondegenerate pairings arise in exterior algebra.
Figure 8.1: Metric Clock
8.4 Hyperplanes and Linear Forms

Actually, Proposition 8.4 below follows from parts (c) and (d) of Theorem 8.1, but we feel that it is also interesting to give a more direct proof.

Proposition 8.4. Let $E$ be a vector space. The following properties hold:

(a) Given any nonnull linear form $f^* \in E^*$, its kernel $H = \text{Ker } f^*$ is a hyperplane.

(b) For any hyperplane $H$ in $E$, there is a (nonnull) linear form $f^* \in E^*$ such that $H = \text{Ker } f^*$.

(c) Given any hyperplane $H$ in $E$ and any (nonnull) linear form $f^* \in E^*$ such that $H = \text{Ker } f^*$, for every linear form $g^* \in E^*$, $H = \text{Ker } g^*$ iff $g^* = \lambda f^*$ for some $\lambda \neq 0$ in $K$.

We leave as an exercise the fact that every subspace $V \neq E$ of a vector space $E$, is the intersection of all hyperplanes that contain $V$.

We now consider the notion of transpose of a linear map and of a matrix.
8.5 Transpose of a Linear Map and of a Matrix

Given a linear map \( f : E \rightarrow F \), it is possible to define a map \( f^\top : F^* \rightarrow E^* \) which has some interesting properties.

**Definition 8.5.** Given a linear map \( f : E \rightarrow F \), the *transpose* \( f^\top : F^* \rightarrow E^* \) of \( f \) is the linear map defined such that

\[
f^\top(v^*) = v^* \circ f,
\]

for every \( v^* \in F^* \).

Equivalently, the linear map \( f^\top : F^* \rightarrow E^* \) is defined such that

\[
\langle v^*, f(u) \rangle = \langle f^\top(v^*), u \rangle,
\]

for all \( u \in E \) and all \( v^* \in F^* \).
It is easy to verify that the following properties hold:

\[
(f + g)^\top = f^\top + g^\top \\
(g \circ f)^\top = f^\top \circ g^\top \\
id_E^\top = \text{id}_{E^*}.
\]

Note the reversal of composition on the right-hand side of \((g \circ f)^\top = f^\top \circ g^\top\).

The equation \((g \circ f)^\top = f^\top \circ g^\top\) implies the following useful proposition.

**Proposition 8.5.** If \(f : E \to F\) is any linear map, then the following properties hold:

(1) If \(f\) is injective, then \(f^\top\) is surjective.

(2) If \(f\) is surjective, then \(f^\top\) is injective.
We also have the following property showing the naturality of the eval map.

**Proposition 8.6.** For any linear map \( f : E \to F \), we have

\[
f^\mathsf{T} \circ \text{eval}_E = \text{eval}_F \circ f,
\]

or equivalently, the following diagram commutes:

\[
\begin{array}{ccc}
E^\ast\ast & \xrightarrow{f^\mathsf{T}} & F^\ast\ast \\
\downarrow{\text{eval}_E} & & \downarrow{\text{eval}_F} \\
E & \xrightarrow{f} & F.
\end{array}
\]
If $E$ and $F$ are finite-dimensional, then $\text{eval}_E$ and $\text{eval}_F$ are isomorphisms, so Proposition 8.6 shows that

$$f^{	op	op} = \text{eval}_F^{-1} \circ f \circ \text{eval}_E.$$  

(*)

The above equation is often interpreted as follows: if we identify $E$ with its bidual $E^{**}$ and $F$ with its bidual $F^{**}$, then $f^{	op	op} = f$.

This is an abuse of notation; the rigorous statement is (*).

**Proposition 8.7.** Given a linear map $f : E \to F$, for any subspace $V$ of $E$, we have

$$f(V)^0 = (f^\top)^{-1}(V^0) = \{w^* \in F^* \mid f^\top(w^*) \in V^0\}.$$  

As a consequence,

$$\text{Ker } f^\top = (\text{Im } f)^0 \quad \text{and} \quad \text{Ker } f = (\text{Im } f^\top)^0.$$
The following theorem shows the relationship between the rank of \( f \) and the rank of \( f^\top \).

**Theorem 8.8.** Given a linear map \( f : E \to F \), the following properties hold.

(a) The dual \((\text{Im } f)^*\) of \( \text{Im } f \) is isomorphic to \( \text{Im } f^\top = f^\top (F^*) \); that is,

\[
(\text{Im } f)^* \cong \text{Im } f^\top.
\]

(b) If \( F \) is finite dimensional, then \( \text{rk}(f) = \text{rk}(f^\top) \).

The following proposition can be shown, but it requires a generalization of the duality theorem.

**Proposition 8.9.** If \( f : E \to F \) is any linear map, then the following identities hold:

\[
\begin{align*}
\text{Im } f^\top &= (\text{Ker } f)^0 \\
\text{Ker } (f^\top) &= (\text{Im } f)^0 \\
\text{Im } f &= (\text{Ker } (f^\top))^0 \\
\text{Ker } (f) &= (\text{Im } f^\top)^0.
\end{align*}
\]
The following proposition shows the relationship between the matrix representing a linear map \( f : E \rightarrow F \) and the matrix representing its transpose \( f^\top : F^* \rightarrow E^* \).

**Proposition 8.10.** Let \( E \) and \( F \) be two vector spaces, and let \((u_1, \ldots, u_n)\) be a basis for \( E \), and \((v_1, \ldots, v_m)\) be a basis for \( F \). Given any linear map \( f : E \rightarrow F \), if \( M(f) \) is the \( m \times n \)-matrix representing \( f \) w.r.t. the bases \((u_1, \ldots, u_n)\) and \((v_1, \ldots, v_m)\), the \( n \times m \)-matrix \( M(f^\top) \) representing \( f^\top : F^* \rightarrow E^* \) w.r.t. the dual bases \((v^*_1, \ldots, v^*_m)\) and \((u^*_1, \ldots, u^*_n)\) is the transpose \( M(f)^\top \) of \( M(f) \).

We now can give a very short proof of the fact that the rank of a matrix is equal to the rank of its transpose.

**Proposition 8.11.** Given a \( m \times n \) matrix \( A \) over a field \( K \), we have \( \text{rk}(A) = \text{rk}(A^\top) \).

Thus, given an \( m \times n \)-matrix \( A \), the maximum number of linearly independent columns is equal to the maximum number of linearly independent rows.
Proposition 8.12. Given any \( m \times n \) matrix \( A \) over a field \( K \) (typically \( K = \mathbb{R} \) or \( K = \mathbb{C} \)), the rank of \( A \) is the maximum natural number \( r \) such that there is an invertible \( r \times r \) submatrix of \( A \) obtained by selecting \( r \) rows and \( r \) columns of \( A \).

For example, the \( 3 \times 2 \) matrix

\[
A = \begin{pmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22} \\
a_{31} & a_{32}
\end{pmatrix}
\]

has rank 2 iff one of the three \( 2 \times 2 \) matrices

\[
\begin{pmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{pmatrix}, \quad \begin{pmatrix}
a_{11} & a_{12} \\
a_{31} & a_{32}
\end{pmatrix}, \quad \begin{pmatrix}
a_{21} & a_{22} \\
a_{31} & a_{32}
\end{pmatrix}
\]

is invertible. We will see in Chapter 4 that this is equivalent to the fact the determinant of one of the above matrices is nonzero.

This is not a very efficient way of finding the rank of a matrix. We will see that there are better ways using various decompositions such as LU, QR, or SVD.
"THE BEAUTY OF THIS IS THAT IT IS ONLY OF THEORETICAL IMPORTANCE, AND THERE IS NO WAY IT CAN BE OF ANY PRACTICAL USE WHATSOEVER."

Figure 8.2: Beauty
8.6 The Four Fundamental Subspaces

Given a linear map \( f : E \to F \) (where \( E \) and \( F \) are finite-dimensional), Proposition 8.7 revealed that the four spaces

\[ \text{Im} f, \text{Im} f^\top, \text{Ker} f, \text{Ker} f^\top \]

play a special role. They are often called the \textit{fundamental subspaces} associated with \( f \).

These spaces are related in an intimate manner, since Proposition 8.7 shows that

\[ \text{Ker} f = (\text{Im} f^\top)^0 \]
\[ \text{Ker} f^\top = (\text{Im} f)^0, \]

and Theorem 8.8 shows that

\[ \text{rk}(f) = \text{rk}(f^\top). \]
It is instructive to translate these relations in terms of matrices (actually, certain linear algebra books make a big deal about this!).

If \( \dim(E) = n \) and \( \dim(F) = m \), given any basis \((u_1, \ldots, u_n)\) of \(E\) and a basis \((v_1, \ldots, v_m)\) of \(F\), we know that \(f\) is represented by an \(m \times n\) matrix \(A = (a_{ij})\), where the \(j\)th column of \(A\) is equal to \(f(u_j)\) over the basis \((v_1, \ldots, v_m)\).

Furthermore, the transpose map \(f^\top\) is represented by the \(n \times m\) matrix \(A^\top\) (with respect to the dual bases).

Consequently, the four fundamental spaces

\[
\text{Im } f, \text{ Im } f^\top, \text{ Ker } f, \text{ Ker } f^\top
\]

correspond to
(1) The \textit{column space} of $A$, denoted by $\text{Im} A$ or $\mathcal{R}(A)$; this is the subspace of $\mathbb{R}^m$ spanned by the columns of $A$, which corresponds to the image $\text{Im} f$ of $f$.

(2) The \textit{kernel} or \textit{nullspace} of $A$, denoted by $\text{Ker} A$ or $\mathcal{N}(A)$; this is the subspace of $\mathbb{R}^n$ consisting of all vectors $x \in \mathbb{R}^n$ such that $Ax = 0$.

(3) The \textit{row space} of $A$, denoted by $\text{Im} A^\top$ or $\mathcal{R}(A^\top)$; this is the subspace of $\mathbb{R}^n$ spanned by the rows of $A$, or equivalently, spanned by the columns of $A^\top$, which corresponds to the image $\text{Im} f^\top$ of $f^\top$.

(4) The \textit{left kernel} or \textit{left nullspace} of $A$ denoted by $\text{Ker} A^\top$ or $\mathcal{N}(A^\top)$; this is the kernel (nullspace) of $A^\top$, the subspace of $\mathbb{R}^m$ consisting of all vectors $y \in \mathbb{R}^m$ such that $A^\top y = 0$, or equivalently, $y^\top A = 0$.

Recall that the dimension $r$ of $\text{Im} f$, which is also equal to the dimension of the column space $\text{Im} A = \mathcal{R}(A)$, is the \textit{rank} of $A$ (and $f$).
Then, some our previous results can be reformulated as follows:

1. The column space $\mathcal{R}(A)$ of $A$ has dimension $r$.
2. The nullspace $\mathcal{N}(A)$ of $A$ has dimension $n - r$.
3. The row space $\mathcal{R}(A^\top)$ has dimension $r$.
4. The left nullspace $\mathcal{N}(A^\top)$ of $A$ has dimension $m - r$.

The above statements constitute what Strang calls the *Fundamental Theorem of Linear Algebra, Part I* (see Strang [32]).
The two statements
\[ \text{Ker } f = (\text{Im } f^\top)^0 \]
\[ \text{Ker } f^\top = (\text{Im } f)^0 \]
translate to

(1) The nullspace of \( A \) is the orthogonal of the row space of \( A \).

(2) The left nullspace of \( A \) is the orthogonal of the column space of \( A \).

The above statements constitute what Strang calls the \textit{Fundamental Theorem of Linear Algebra, Part II} (see Strang [32]).

Since vectors are represented by column vectors and linear forms by row vectors (over a basis in \( E \) or \( F' \)), a vector \( x \in \mathbb{R}^n \) is orthogonal to a linear form \( y \) if
\[ yx = 0. \]
Then, a vector $x \in \mathbb{R}^n$ is orthogonal to the row space of $A$ iff $x$ is orthogonal to every row of $A$, namely $Ax = 0$, which is equivalent to the fact that $x$ belong to the nullspace of $A$.

Similarly, the column vector $y \in \mathbb{R}^m$ (representing a linear form over the dual basis of $F^*$) belongs to the nullspace of $A^\top$ iff $A^\top y = 0$, iff $y^\top A = 0$, which means that the linear form given by $y^\top$ (over the basis in $F$) is orthogonal to the column space of $A$.

Since (2) is equivalent to the fact that the column space of $A$ is equal to the orthogonal of the left nullspace of $A$, we get the following criterion for the solvability of an equation of the form $Ax = b$:

The equation $Ax = b$ has a solution iff for all $y \in \mathbb{R}^m$, if $A^\top y = 0$, then $y^\top b = 0$. 
Indeed, the condition on the right-hand side says that $b$ is orthogonal to the left nullspace of $A$, that is, that $b$ belongs to the column space of $A$.

This criterion can be cheaper to check that checking directly that $b$ is spanned by the columns of $A$. For example, if we consider the system

$$
\begin{align*}
    x_1 - x_2 &= b_1 \\
    x_2 - x_3 &= b_2 \\
    x_3 - x_1 &= b_3
\end{align*}
$$

which, in matrix form, is written $Ax = b$ as below:

$$
\begin{pmatrix}
    1 & -1 & 0 \\
    0 & 1 & -1 \\
    -1 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2 \\
    x_3
\end{pmatrix} =
\begin{pmatrix}
    b_1 \\
    b_2 \\
    b_3
\end{pmatrix},
$$

we see that the rows of the matrix $A$ add up to 0.
In fact, it is easy to convince ourselves that the left nullspace of $A$ is spanned by $y = (1, 1, 1)$, and so the system is solvable iff $y^\top b = 0$, namely

$$b_1 + b_2 + b_3 = 0.$$  

Note that the above criterion can also be stated negatively as follows:

The equation $Ax = b$ has no solution iff there is some $y \in \mathbb{R}^m$ such that $A^\top y = 0$ and $y^\top b \neq 0$. 
Figure 8.3: Brain Size?