UNIVERSITY OF CALIFORNIA, DAVIS
Department of Electrical and Computer Engineering

EEC250 | Linear Systems and Signals | Fall 2009

Lecture 3

Topics:

a) Computation and characterization of the four fundamental spaces of a matrix

b) Linear map, change of basis, similarity transformation

As we saw in Lecture 1, we can associate four fundamental spaces to an arbitrary $m \times n$ matrix A:

(i) The column space

$$\mathcal{R}(A) = \{\mathbf{b} \in \mathbb{R}^m :, \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n\} \tag{1}$$

(ii) The row space

$$\begin{aligned} \mathcal{R}(A^T) &= \{\mathbf{c} \in \mathbb{R}^n : \mathbf{c} = A^T\mathbf{y} \text{ for some } \mathbf{y} \in \mathbb{R}^m\} \\ &= \{\mathbf{c} \in \mathbb{R}^n : \mathbf{c}^T = \mathbf{y}^T A \text{ for some } \mathbf{y} \in \mathbb{R}^m\} \end{aligned} \tag{2}$$

(iii) The right null space

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\} \tag{3}$$

(iv) The left null space

$$\mathcal{N}(A^T) = \{\mathbf{y} \in \mathbb{R}^m : A^T\mathbf{y} = \mathbf{0}\} = \{\mathbf{y} \in \mathbb{R}^m : \mathbf{y}^T A = \mathbf{0}^T\}\,. \tag{4}$$

It turns out that all four fundamental spaces can be computed from the row echelon form E (or the reduced row echelon form E_r) of the matrix A, and from the matrix M (resp. M_r) corresponding to the combination of the elementary operations bringing A to the form E (resp. E_r). To see this, consider the following example.

Example: Let

$$A = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 4 & 6 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

Then, we have

$$MA = E$$

with

$$M = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \quad , \quad E = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since elementary row operations do not change the row space of A, the row space of A is given by the nonzero rows of E. Thus, we have:

$$\textit{row space} \ = \mathcal{R}(A^T) = \ \text{space spanned by } \left\{ \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} , \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

The elementary row operations change the column space of A, but if a set of columns of A is linearly independent, the same set of columns of E is independent. This is due to the fact that the solutions of $A\mathbf{x} = \mathbf{0}$ do not change if we we perform elementary operations on the rows of A Then,. since the columns of E where the pivot elements are located (here, the second and fourth columns) constitute a basis for the column space of E, the corresponding columns of A constitute a basis for the column space of A. Thus,

$$\textit{column space} \ = \mathcal{R}(A) = \ \text{space spanned by } \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} , \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix} \right\},$$

where the two vectors spanning $\mathcal{R}(A)$ are the second and fourth columns of A.

The left null space of A is obtained by examining the rows of E that are zero. Since these rows are zero, the corresponding row vectors of M are in the left null space of A. These row vectors are necessarily independent since M is invertible, so that they constitute a basis of the left null space of A. Here the third row of $E = 0$, so

$$\textit{left null space} \ = \mathcal{N}(A^T) = \ \text{space spanned by } \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\},$$

where the vector spanning $\mathcal{N}(A^T)$ is the transpose of the third row of M.

To find the right null space of A, it is convenient to compute the reduced row echelon form E_r. For the example considered here

$$M_r A = E_r = \begin{bmatrix} 0 & 1 & 2 & 0 & -2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

with

$$M_r = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} M = \begin{bmatrix} 4 & -3 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}.$$

A basis of the right null space of A is obtained by expressing each *nonpivot* column of E_r (here, the first, third and fifth columns) in terms of the pivot columns. Since the pivot columns of E_r are the standard orthonormal basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_r$, where r = rank of A, this is easy to accomplish. Here

$$1^{\text{st}} \text{ column of } E_r = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0. \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0. \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$3^{\text{rd}} \text{ column of } E_r = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = 2. \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0. \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$5^{\text{th}} \text{ column of } E_r = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix} = -2. \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 2. \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

The above expressions yield linear dependence relations between the columns of E_r and therefore of A, so that

$$\textit{right null space } = \mathcal{N}(A) = \text{ space spanned by } \{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \\ -2 \\ 1 \end{bmatrix} \}.$$

$\square$

By using the procedure described above, we see that if A has rank r, we can generate $n - r$ vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{n-r}$ (one for each nonbasic column of E_r) in the null space of A. These vectors are clearly independent since each has one nonbasic variable equal to one, and all other nonbasic variables equal to zero. To see that they span the whole right null space $\mathcal{N}(A)$, let $\mathbf{y}$ be an arbitrary vector of $\mathcal{N}(A)$, so that $A\mathbf{y} = 0$. In this case we have also $E_r\mathbf{y} = 0$. Then, consider

$$\mathbf{z} = \mathbf{y} - \sum_{k=1}^{n-r} y_{i_k} \mathbf{x}_k \tag{5}$$

where $\mathbf{x}_k$, $1 \leq k \leq n - r$ are the vectors of $\mathcal{N}(A)$ constructed above and i_k is the index of the k-th nonbasic variable of $\mathbf{y}$, and y_{i_k} the corresponding variable. We have

$$E_r \mathbf{z} = 0 \tag{6}$$

where all the nonbasic variables of $\mathbf{z}$ are zero. Since the submatrix of E_r obtained by retaining only the pivot columns of E_r has the form

$$\begin{bmatrix} I_r \\ 0 \end{bmatrix},$$

where I_r denotes the $r \times r$ identity matrix, the relation (6) implies $\mathbf{z} = \mathbf{0}$. Thus

$$\mathbf{y} = \sum_{k=1}^{n-r} y_{i_k} \mathbf{x}_k \,, \tag{7}$$

so that $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-r}\}$ is a basis of $\mathcal{N}(A)$, as claimed.

To explore further the structure of the right and left null spaces of A, it is convenient to introduce the following concept.

Orthogonal complement: Let W be a subspace of $\mathbb{R}^n$. Its orthogonal complement $W^\perp$ is given by

$$W^\perp = \{\mathbf{u} \in \mathbb{R}^n \;:\; \mathbf{u}^T\mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\} \,. \tag{8}$$

To verify that $W^\perp$ is a subspace of $\mathbb{R}^n$, let $\mathbf{u}_1$ and $\mathbf{u}_2$ be two vectors of $W^\perp$, so that

$$\mathbf{u}_1^T\mathbf{w} = \mathbf{u}_2^T\mathbf{w} = 0$$

for all $\mathbf{w} \in W$. Then if a and b are two arbitrary scalars of $\mathbb{R}$, we have

$$(a\mathbf{u}_1 + b\mathbf{u}_2)^T\mathbf{w} = 0$$

for all $\mathbf{w} \in W$, so that $a\mathbf{u}_1 + b\mathbf{u}_2 \in W^\perp$.

Example: Let W be the one-dimensional subspace of $\mathbb{R}^3$ formed by vectors colinear with

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} .$$

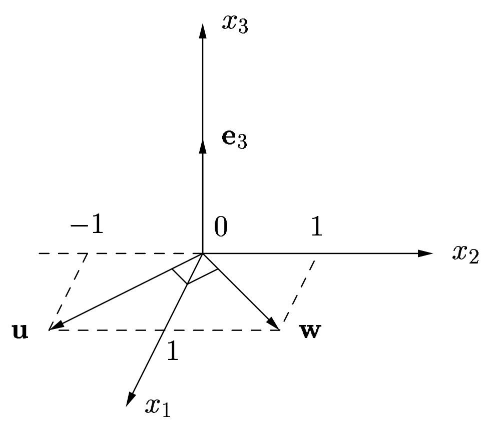

Figure 1: Orthononal complement of the space W spanned by $\mathbf{w}$.

Then $W^\perp$ is the plane perpendicular to the vector $\mathbf{w}$. It is spanned by the vectors

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

as shown in Fig. 1. □

Properties of orthogonal complements:

(i) No nonzero vector can be perpendicular to itself:

$$W \cap W^{\perp} = \{\mathbf{0}\} .$$

(ii) The orthogonal complement of the orthogonal complement is the original subspace:

$$(W^{\perp})^{\perp} = W .$$

This can be seen by noting that the vectors $\mathbf{w}$ of W have the property that $\mathbf{w}^T\mathbf{u} = 0$ for all vectors $\mathbf{u}$ of $W^{\perp}$.

(iii) $\mathbb{R}^n$ is the direct sum of W and $W^{\perp}$:

$$\mathbb{R}^n = W \oplus W^{\perp}$$

This means that any vector $\mathbf{x}$ of $\mathbb{R}^n$ can be decomposed as

$$\mathbf{x} = \mathbf{x}^{||} + \mathbf{x}^{\perp}$$

where $\mathbf{x}^{||} \in W$ and $\mathbf{x}^{\perp} \in W^{\perp}$.

Property (iii) can be verified by observing that an arbitrary basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_q\}$ of W can always be completed into a basis of $\mathbb{R}^n$ by selecting basis vectors $\{\mathbf{e}_{q+1}, \ldots, \mathbf{e}_n\}$ which are perpendicular to W. Then an arbitrary vector $\mathbf{x}$ of $\mathbb{R}^n$ can be expressed as

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i = \mathbf{x}^{||} + \mathbf{x}^{\perp} ,$$

where

$$\begin{aligned} \mathbf{x}^{||} &= \sum_{i=1}^{q} x_i \mathbf{e}_i \in W \\ \mathbf{x}^{\perp} &= \sum_{i=q+1}^{n} x_i \mathbf{e}_i \in W^{\perp} . \end{aligned}$$

The four fundamental spaces of A can be characterized as follows.

Theorem: We have

$$\mathcal{N}(A) = \mathcal{R}(A^T)^{\perp} \tag{9}$$

$$\mathcal{N}(A^T) = \mathcal{R}(A)^{\perp} . \tag{10}$$

Proof: Since (10) is obtained by replacing A by A^T in (9), we only need to prove (7). If $\mathbf{x}$ is an arbitrary vector of $\mathcal{N}(A)$, it satisfies

$$A\mathbf{x} = \mathbf{0} .$$

Similarly, if $\mathbf{c}$ is an arbitrary vector of $\mathcal{R}(A^T)$, we have

$$\mathbf{u}^T A = \mathbf{c}^T \tag{11}$$

for some $\mathbf{u} \in \mathbb{R}^m$. Multiplying (11) on the right by $\mathbf{x}$ yields

$$\mathbf{c}^T \mathbf{x} = \mathbf{u}^T A \mathbf{x} = 0$$

for $\mathbf{x} \in \mathcal{N}(A)$ and $\mathbf{c} \in \mathcal{R}(A^T)$, so that $\mathcal{N}(A)$ and $\mathcal{R}(A^T)$ are two mutually orthogonal subspaces of $\mathbb{R}^n$.

In order to show that $\mathcal{N}(A) = \mathcal{R}(A^T)^\perp$, we must prove that *all* the vectors orthogonal to $\mathcal{R}(A^T)$ are in $\mathcal{N}(A)$. But the dimension of $\mathcal{R}(A^T)$, the row space of A is $r = \text{rank}(A)$. This implies that its orthogonal complement has dimension $n - r$. The basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_{n-r}\}$ of $N(A)$ that we have constructed earlier has precisely $n - r$ elements (one for each nonbasic column of E_r), so that $\mathcal{N}(A) = \mathcal{R}(A^T)^\perp$. □

From the above result we see that

$$\begin{aligned} \mathbb{R}^n &= \mathcal{R}(A^T) \oplus \mathcal{N}(A) & (12) \\ \mathbb{R}^m &= \mathcal{R}(A) \oplus \mathcal{N}(A^T)\,. & (13) \end{aligned}$$

To interpret these decompositions, note that A just maps its domain $\mathbb{R}^n$ into its range space $\mathcal{R}(A) \subset \mathbb{R}^m$. Then, since A maps vectors in its null space $\mathcal{N}(A)$ onto the zero vector $\mathbf{0}$, the decomposition (12) shows that the range space $\mathcal{R}(A)$ (the column space of A) is effectively generated by applying A to $\mathcal{R}(A^T)$ (the row space of A). In other words, *A maps its row space into its column space*, as indicated by the diagram shown in Fig. 2.

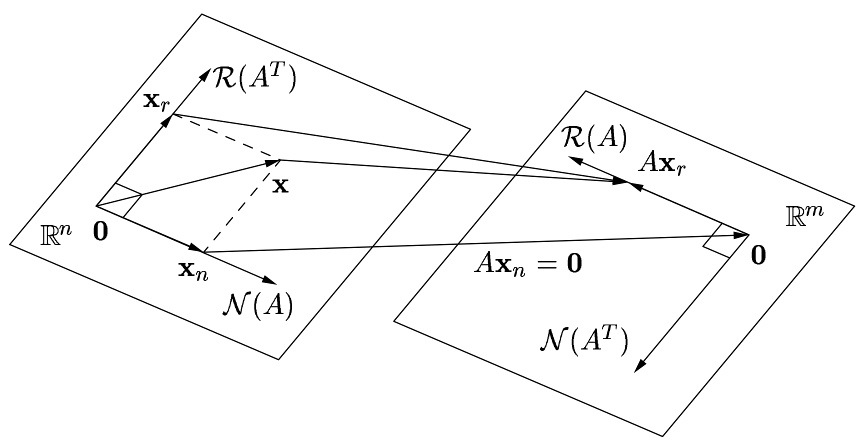

Figure 2: Mapping of the domain $\mathbb{R}^n$ of A into its range space.

By duality, A^T is a map of $I\!R^m$ into $R(A^T) \subset I\!R^n$. It maps the column space of A into its row space.

Another byproduct of decompositions (12) and (13) is that if A is a matrix of size $m \times n$ and rank r, we have

$$
\begin{aligned}
n = \dim\,(\mathbb{R}^n) &= \dim \text{ domain } (A) \\
&= \dim \mathcal{R}(A) + \dim \mathcal{N}(A) = r + (n - r) \qquad (14) \\
m = \dim\,(\mathbb{R}^m) &= \dim \text{ domain } (A^T) \\
&= \dim \mathcal{R}(A^T) + \dim \mathcal{N}(A^T) = r + (m - r)\,, \qquad (15)
\end{aligned}
$$

where $\dim\,(V)$ denotes the dimension of a vector space V, and where we have used the fact that

$$\dim \mathcal{R}(A) = \dim \mathcal{R}(A^T) = r \qquad (16)$$

(the dimension of the row or column space of A equals its rank). To interpret (14), note that A acts on vectors $\mathbf{x} \in \mathbb{R}^n$. These vectors have n degrees of freedom. They are mapped by A into vectors $\mathbf{y} = A\mathbf{x}$ in the range of A which have only r degrees of freedom. The remaining $n - r$ degrees of freedom can be accounted for by noting that the null space $\mathcal{N}(A)$ of vectors $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$ has dimension $n - r$.

Summary: If A has size $m \times n$ and rank r, we have

(i) $\mathcal{N}(A) = \mathcal{R}(A^T)^{\perp}$, $\mathcal{N}(A^T) = \mathcal{R}(A)^{\perp}$

(ii) $\dim \mathcal{R}(A) = \dim \mathcal{R}(A^T) = r$
$\dim \mathcal{N}(A) = n - r$, $\dim N(A^T) = m - r$.

Linear map: Consider two vector spaces V and W and a map $L : \mathbf{v} \in V \to \mathbf{w} = L\mathbf{v} \in W$. L is a linear map if when $\mathbf{w}_1 = L\mathbf{v}_1$ and $\mathbf{w}_2 = L\mathbf{v}_2$, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are arbitrary vectors of V, then

$$a\mathbf{w}_1 + b\mathbf{w}_2 = L(a\mathbf{v}_1 + b\mathbf{v}_2)$$

for arbitrary real scalars a and b.

Example: Consider the discrete-time linear time-invariant system with impulse response $h(k)$ shown in Fig. 3.

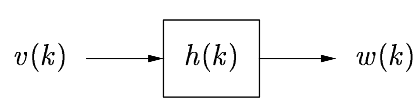

Figure 3: Discrete-time linear time-invariant system.

The output sequence $w(k)$ corresponding to an input sequence $v(k)$ is given by

$$w(k) = (h * v)(k) = \sum_{l=-\infty}^{\infty} h(k - l)v(l)\,. \qquad (17)$$

If the impulse response $h(\cdot)$ is summable, so that the system is BIBO stable, the system (17) defines a linear map from the vector space $V = l^\infty(\mathbb{Z})$ of bounded integer-indexed sequences into $W = l^\infty(\mathbb{Z})$. It is easy to verify that this map is linear since

$$\begin{aligned} h * (av_1 + bv_2) &= ah * v_1 + bh * v_2 \\ &= aw_1 + bw_2 \,, \end{aligned}$$

where $w_i = h * v_i$ for $i = 1, 2$.

Matrix representation of a linear map: We restrict our attention to the case where the vector spaces v and W are finite dimensional. Let $\{\mathbf{e}_j,\ 1 \leq j \leq n\}$ be a basis of V and $\{\mathbf{f}_i,\ 1 \leq i \leq m\}$ a basis of W. Then, arbitrary vectors $\mathbf{v} \in V$ and $\mathbf{w} \in W$ can be expressed in terms of the basis vectors $\{\mathbf{e}_j,\ 1 \leq j \leq n\}$ and $\{\mathbf{f}_i,\ 1 \leq i \leq m\}$ as

$$\mathbf{v} = \sum_{j=1}^{n} x_j \mathbf{e}_j \quad , \quad \mathbf{w} = \sum_{i=1}^{m} y_i \mathbf{f}_i \,.$$

If L is a linear map, then

$$\mathbf{w} = L\mathbf{v} = \sum_{j=1}^{n} x_j L\mathbf{e}_j$$

where each vector $L\mathbf{e}_j \in W$ can be expressed in terms of the basis $\{\mathbf{f}_i,\ 1 \leq i \leq m\}$ of W as

$$L\mathbf{e}_j = \sum_{i=1}^{m} a_{ij} \mathbf{f}_i \,, \tag{18}$$

for $1 \leq j \leq n$. If $\mathbf{w} = L\mathbf{v}$, we have therefore

$$\mathbf{w} = \begin{bmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_m \end{bmatrix} \mathbf{y} = \begin{bmatrix} L\mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix} \mathbf{x} \tag{19}$$

with

$$\mathbf{y} \triangleq \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \quad , \quad \mathbf{x} \triangleq \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

and

$$\begin{bmatrix} L\mathbf{e}_1 & \dots & L\mathbf{e}_n \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_m \end{bmatrix} A \tag{20}$$

Substituting (20) inside (19) yields

$$\begin{bmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_m \end{bmatrix} (\mathbf{y} - A\mathbf{x}) = \mathbf{0} \,, \tag{21}$$

and since the vectors $\{\underline{f}_1, \dots, \underline{f}_m\}$ are linearly independent, this implies

$$\mathbf{y} = A\mathbf{x} \,. \tag{22}$$

Thus, the matrix $A = (a_{ij},\ 1 \leq i \leq m,\ 1 \leq j \leq n)$ obtained by expressing the transformed basis vectors $L\mathbf{e}_j$ of V in terms of the basis vectors of W in (18), expresses the coordinates $\mathbf{y}$ of $\mathbf{w} = L\mathbf{v}$ in terms of the coordinates $\mathbf{x}$ of $\mathbf{v}$.

Example: Let $V = W = \mathbb{R}^{@}$. Then a basis for both V and W is given by

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

Let L be a $\pi/4$ rotation about the origin whose effect on an arbitrary vector $\mathbf{v}$ is shown in figrota.

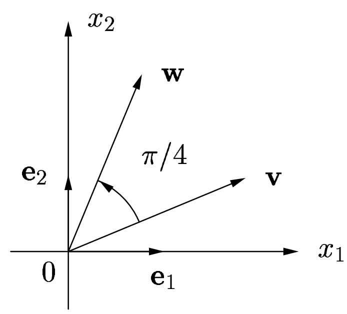

Figure 4: $\pi/4$ rotation about the origin in $\mathbb{R}^2$.

L is clearly linear and

$$\begin{aligned} L\mathbf{e}_1 &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\mathbf{e}_1 + \frac{1}{\sqrt{2}}\mathbf{e}_2 \\ L\mathbf{e}_2 &= \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -\frac{1}{\sqrt{2}}\mathbf{e}_1 + \frac{1}{\sqrt{2}}\mathbf{e}_2 , \end{aligned}$$

so that the matrix representation of L is

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} .$$

Change of basis: Consider two bases of $\mathbb{R}^n$, say $\{\mathbf{e}_i^{\text{old}}, 1 \le i \le n\}$ and $\{\mathbf{e}_i^{\text{new}}, 1 \le i \le n\}$. Since both are bases we can express any vector $\mathbf{e}_j^{\text{new}}$ of the new basis in terms of the old basis vectors, i.e.

$$\mathbf{e}_j^{\text{new}} = \sum_{i=1}^{n} \mathbf{e}_i^{\text{old}} t_{ij} \tag{23}$$

Let

$$E_{\text{new}} \triangleq \begin{bmatrix} \mathbf{e}_1^{\text{new}} & \dots & \mathbf{e}_n^{\text{new}} \end{bmatrix} \quad E_{\text{old}} \triangleq \begin{bmatrix} \mathbf{e}_1^{\text{old}} & \dots & \mathbf{e}_n^{\text{old}} \end{bmatrix}$$

denote the two square $n \times n$ matrices whose columns correspond to the new and old basis vectors, respectively. Then identity (21) can be rewritten in matrix form as

$$E_{\text{new}} = E_{\text{old}} T , \tag{24}$$

with $T = (t_{ij},\ 1 \leq i,\ j \leq n$. Similarly, we can express the old basis vectors in terms of the new basis vectors as

$$\mathbf{e}_j^{old} = \sum_{i=1}^{n} \mathbf{e}_i^{new} s_{ij}$$

for $1 \leq j \leq n$, or equivalently as

$$E_{old} = E_{new} S \tag{25}$$

with $S = (s_{ij}) \quad 1 \leq i,\ j \leq n$. Substituting (25) inside (24) yields

$$E_{new}(I - ST) = 0\,. \tag{26}$$

But since the columns of E_{new} are linearly independent, E_{new} is invertible so that

$$I = ST = TS\,, \tag{27}$$

i.e. $S = T^{-1}$.

In terms of both the old and new bases, an arbitrary vector $\mathbf{v}$ of $\mathbb{R}^n$ can be expressed as

$$\begin{aligned}\mathbf{v} &= \sum_{i=1}^{n} x_i^{\text{old}} \mathbf{e}_i^{\text{old}} = \sum_{i=1}^{n} x_i^{\text{new}} \mathbf{e}_i^{\text{new}} \\ &= E_{\text{old}} \mathbf{x}_{\text{old}} = E_{\text{new}} \mathbf{x}_{\text{new}} \end{aligned} \tag{28}$$

with

$$\mathbf{x}_{\text{old}} \triangleq \begin{bmatrix} x_1^{\text{old}} \\ \vdots \\ x_n^{\text{old}} \end{bmatrix} \qquad \mathbf{x}_{\text{new}} \triangleq \begin{bmatrix} x_1^{\text{new}} \\ \vdots \\ x_n^{\text{new}} \end{bmatrix}.$$

But $E_{\text{old}} = E_{\text{new}} T^{-1}$, so that

$$\underline{x}_{\text{new}} = T^{-1} \underline{x}_{old}\,. \tag{29}$$

Comparing (25) and (29), we see that basis vectors and vector coordinates are transformed in *opposite ways*, i.e. going from the old to the new basis vectors corresponds to a multiplication by T, but going from the old to the new coordinates corresponds to a multiplication by T^{-1}.

Now let L be a linear map from $\mathbb{R}^n$ to $\mathbb{R}^n$ which is represented by the matrix A_{old} with respect to the basis $\{\mathbf{e}_i^{\text{old}},\ 1 \leq i \leq n\}$. In other words, if $\mathbf{w} = L\mathbf{v}$, the coordinates $\mathbf{y}_{\text{old}}$ and $\mathbf{x}_{\text{old}}$ of $\mathbf{w}$ and $\mathbf{v}$ are related through

$$\mathbf{y}_{\text{old}} = A_{\text{old}} \mathbf{y}_{\text{old}}\,.$$

If we now perform the change of basis (24), we find

$$\mathbf{y}_{\text{new}} = T^{-1} \mathbf{y}_{\text{old}} = T^{-1} A_{\text{old}} T \mathbf{x}_{\text{new}}\,,$$

so that with respect to the new basis $\{\mathbf{e}_i^{\text{new}},\ 1 \leq i \leq n\}$, L is represented by

$$A_{\text{new}} = T^{-1} A_{\text{old}} T\,. \tag{30}$$

The transformation (30) relating A_{new} and A_{old} is called a *similarity transformation.*

Example: Consider the change of basis of $\mathbb{R}^2$ where we go from the standard orthonormal basis

$$\mathbf{e}_1^{\text{old}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{e}_2^{\text{old}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

to the new basis

$$\mathbf{e}_1^{\text{new}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \mathbf{e}_2^{\text{new}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

as shown in Fig. 5 below.

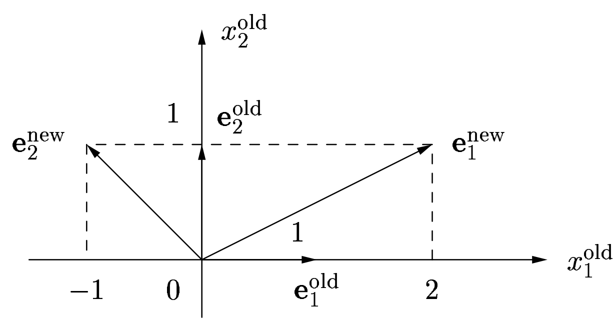

Figure 5: Basis vectors of the old and new bases.

Clearly

$$\begin{bmatrix} \mathbf{e}_1^{\text{new}} & \mathbf{e}_2^{\text{new}} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1^{\text{old}} & \mathbf{e}_2^{\text{old}} \end{bmatrix} T$$

with

$$T = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix},$$

and

$$\begin{bmatrix} \mathbf{e}_1^{\text{old}} & \mathbf{e}_2^{\text{old}} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1^{\text{new}} & \mathbf{e}_2^{\text{new}} \end{bmatrix} T^{-1}.$$

Then if $\mathbf{x}$ is an arbitrary vector of $\mathbb{R}^2$ whose coordinates with respect to $\{\mathbf{e}_1^{\text{old}}, \mathbf{e}_2^{\text{old}}\}$ are $(x_1^{\text{old}}, x_2^{\text{old}})$, its coordinates with respect to $\{\mathbf{e}_1^{\text{new}}, \mathbf{e}_2^{\text{new}}\}$ are given by

$$\begin{bmatrix} x_1^{\text{new}} \\ x_2^{\text{new}} \end{bmatrix} = T^{-1} \begin{bmatrix} x_1^{\text{old}} \\ x_2^{\text{old}} \end{bmatrix}$$

with

$$T^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}.$$

Now let R be a rotation by $\pi/2$ with respect to the origin. R is a linear map which maps $\mathbf{e}_1^{\text{old}}$ into $\mathbf{e}_2^{\text{old}}$ and $\mathbf{e}_2^{\text{old}}$ into $-\mathbf{e}_1^{\text{old}}$, so that

$$\begin{bmatrix} R\mathbf{e}_1^{\text{old}} & R\mathbf{e}_2^{\text{old}} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1^{\text{old}} & \mathbf{e}_2^{\text{old}} \end{bmatrix} A_{\text{old}}$$

where

$$A_{\text{old}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

is the matrix representation of the rotation R in the basis $\{\mathbf{e}_1^{\text{old}}, \mathbf{e}_2^{\text{old}}\}$. In the new basis, R is represented by

$$\begin{aligned} A_{\text{new}} &= T^{-1} A_{\text{old}} T \\ &= \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 1 & -2 \\ 5 & 1 \end{bmatrix}. \end{aligned}$$