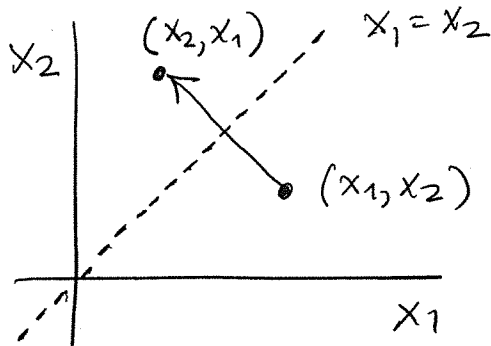


Sec. 3.7 Linear transformations

① Examples and Definition

Ex. 1

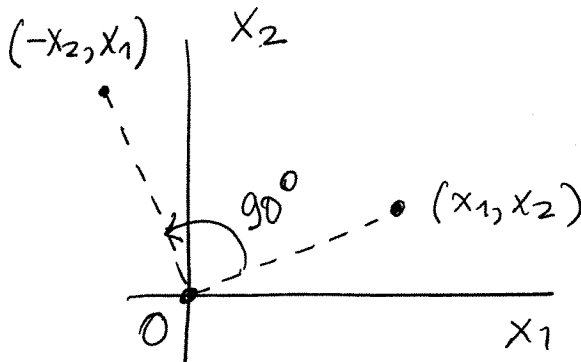


Reflection about $x_1 = x_2$:

$$T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$T \left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right] = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

Ex. 2



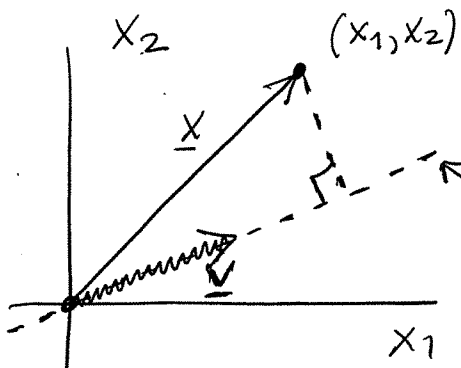
Rotation about origin by 90° :

$$T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$T \left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right] = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

(For a rotation by θ , the transformation is found on p. 236 of textbooks; see also Project 1.)

Ex. 3



Projection on direction of \underline{v}

$$T: \mathbb{R}^2 \rightarrow \mathbb{R}^1$$

$$T \left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right] = P_{\underline{v}}(\underline{x})$$

$$= \frac{\underline{v}^T \underline{x}}{\|\underline{v}\|^2} \cdot \underline{v}$$

sec. 3,6

Def Let V and W be subspaces of \mathbb{R}^n .
 $T: V \rightarrow W$ is a **linear transformation (l.t.)**

iff for any $\underline{x}, \underline{y}$ in V it satisfies:

(a) $T(\underline{x} + \underline{y}) = T(\underline{x}) + T(\underline{y})$

(b) $T(c \cdot \underline{x}) = c \cdot T(\underline{x})$ for any scalar c .

Corollary 1

$(T \text{ is a l.t.}) \Leftrightarrow (T(a\underline{x} + b\underline{y}) = aT(\underline{x}) + bT(\underline{y}))$

Corollary 2 If T is a l.t., then for any vectors $\underline{v}_1, \dots, \underline{v}_p$ and scalars c_1, \dots, c_p :

$T(c_1\underline{v}_1 + \dots + c_p\underline{v}_p) = c_1T(\underline{v}_1) + \dots + c_pT(\underline{v}_p)$

How can we tell if a given T is a l.t.?

We can check properties (a) & (b); see Ex. 1.

However, **a more efficient method will be shown in topic ②**, coming up next.

② Matrix of a l.t.

Revisit Ex. 1 and rewrite the result of T as:

$T\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \equiv \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$
 $= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

↑ some matrix, independent of \underline{x}

Similarly, in Ex. 2: $T\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Thus, in both cases, for all \underline{x} :

$$T(\underline{x}) = A \underline{x} \quad \leftarrow \text{some matrix, independent of } \underline{x}$$

Claim:

$$(T \text{ is l.t.}) \Leftrightarrow (T(\underline{x}) = A \underline{x})$$

Proof (only of " \Leftarrow ")

All we need to do is check that $T(\underline{x}) = A \underline{x}$ satisfies properties (a) and (b).

$$(a) \quad A(\underline{x} + \underline{y}) \stackrel{?}{=} A \underline{x} + A \underline{y}$$

This is true by the distributive property.

$$(b) \quad A(c \cdot \underline{x}) \stackrel{?}{=} c A \underline{x}$$

This is true because one can move a scalar before or after A .

q. e. d.

Thus, by this Claim, the transformations in Ex. 1 & 2 are both l.t.

Ex. 4 Let us show that the projection, considered in Ex. 3, is a l.t. According to the Claim, we need to exhibit matrix A (independent of \underline{x} !) such that $P_{\underline{v}}(\underline{x}) = A \underline{x}$.

Sol'n: 1) We first simplify the formula for $P_{\underline{v}}(\underline{x})$. The motivation for this simplification is as follows. When we say "projection on \underline{v} ", we actually mean "projection on the direction of \underline{v} ". So, the length of \underline{v} (but not the length of \underline{x} !) does not matter, and so what we do below is eliminate the length of \underline{v} from the formula.

$$P_{\underline{v}}(\underline{x}) = \frac{\underline{v}^T \underline{x}}{\|\underline{v}\|^2} \cdot \underline{v} = \left(\frac{\underline{v}^T}{\|\underline{v}\|} \right) \cdot \underline{x} \cdot \left(\frac{\underline{v}}{\|\underline{v}\|} \right) = (\underline{w}^T \underline{x}) \cdot \underline{w},$$

where $\underline{w} = \underline{v} / \|\underline{v}\|$ is the unit vector along \underline{v} .

So,

$$P_{\underline{v}}(\underline{x}) \equiv P_{\underline{w}}(\underline{x}) = (\underline{w}^T \underline{x}) \cdot \underline{w}$$

2) We now transform this into the form $A \cdot \underline{x}$.

$$\underbrace{(\underline{w}^T \underline{x})}_{\text{scalar}} \cdot \underline{w} = \underline{w} \cdot (\underline{w}^T \underline{x}) \stackrel{\text{associative rule}}{=} \underbrace{(\underline{w} \underline{w}^T)}_{\substack{\text{matrix} \\ \text{independent} \\ \text{of } \underline{x}}} \underline{x}$$

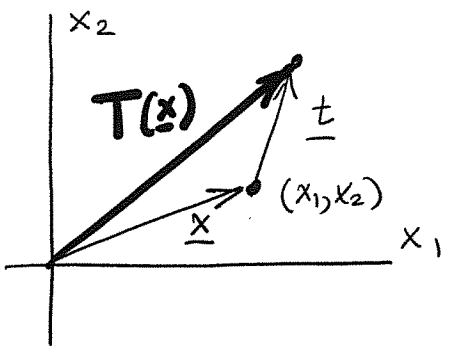
Thus,

$$\left(\begin{array}{l} \text{the matrix of projection} \\ \text{on the direction of } \underline{v} \end{array} \right) = \underline{w} \underline{w}^T$$

where $\underline{w} = \underline{v} / \|\underline{v}\|$, the unit vector along \underline{v} .

And, according to the Claim on p. 15-3, projection is a l.t..

Ex. 5 Translation is not a l. t.



Translation moves (translates) all points in the plane by the same vector (say, \underline{t}).

So: $T(\underline{x}) = \underline{x} + \underline{t}$

(see the figure; recall Calculus).

There is no way to write $\underline{x} + \underline{t} = A \cdot \underline{x}$ that would work for all \underline{x} with an A that is independent of \underline{x} (if you are not convinced, try to find $a = \text{const}$ such that $x + 1 = a \cdot x$ for all x).

③ l. t. and expansion over a basis

Let $\{\underline{v}_1, \dots, \underline{v}_p\}$ be a basis for some subspace W of \mathbb{R}^n , and let \underline{x} be any vector in W . We know that \underline{x} can be expanded over a basis:

$$\underline{x} = c_1 \underline{v}_1 + \dots + c_p \underline{v}_p,$$

where c_1, \dots, c_p are coordinates of \underline{x} in basis $\{\underline{v}_1, \dots, \underline{v}_p\}$.

Then for some l. t. T :

$$T(\underline{x}) = T(c_1 \underline{v}_1 + \dots + c_p \underline{v}_p) \stackrel{\text{Corollary 2 on p. 15-2}}{=} c_1 T(\underline{v}_1) + \dots + c_p T(\underline{v}_p).$$

This implies that we now have

two ways to define the action of a l. t.:

1] As per the Claim on p. 15-3, we can define $T(\underline{x}) = A\underline{x}$ (which is equivalent to what we did in Ex. 1 & 2).

2] If we know a basis $\{v_1, \dots, v_p\}$ in W and know how a l.t. transforms it, i.e. know $\{T(v_1), \dots, T(v_p)\}$, we can find for any \underline{x} :

$$T(\underline{x}) = c_1 T(v_1) + \dots + c_p T(v_p)$$

↑
coordinates of \underline{x} in basis $\{v_1, \dots, v_p\}$

Ex. 6 (see also Ex. 6 in book)

$$\text{Let } T\left[\begin{pmatrix} -1 \\ 0 \end{pmatrix}\right] = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad T\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

Find $T\left[\begin{pmatrix} 8 \\ 9 \end{pmatrix}\right]$.

Sol'n: 1) $\underbrace{\begin{pmatrix} -1 \\ 0 \end{pmatrix}}_{v_1}, \underbrace{\begin{pmatrix} 2 \\ 1 \end{pmatrix}}_{v_2}$ are not proportional,

\Rightarrow not $\parallel \Rightarrow$ they form a basis in \mathbb{R}^2 .

Expand $\underline{x} = \begin{pmatrix} 8 \\ 9 \end{pmatrix}$ over this basis:

$$\begin{pmatrix} 8 \\ 9 \end{pmatrix} = c_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow$$

$$\left(\begin{array}{cc|c} -1 & 2 & 8 \\ 0 & 1 & 9 \end{array} \right) \Rightarrow \begin{array}{l} -c_1 + 2c_2 = 8 \\ c_2 = 9 \end{array} \begin{array}{l} \nearrow \\ \text{by inspection} \end{array} \Rightarrow \begin{array}{l} c_1 = 10 \\ c_2 = 9 \end{array}$$

Note: When a problem like this is given on a quiz or test, you are allowed to find c_1, c_2 not by REF but by inspection. In all other problems you must still use REF.

2) Now we can use the formula:

$$T(\underline{x}) = c_1 T(\underline{v}_1) + c_2 T(\underline{v}_2) = 10 \begin{pmatrix} 3 \\ 4 \end{pmatrix} + 9 \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \underline{\underline{\begin{pmatrix} 75 \\ 94 \end{pmatrix}}}.$$

The two ways, 1 and 2, to define T must be equivalent. We can easily "go from 1 to 2": if we know the A , we can find

$$\{T(\underline{v}_1), \dots, T(\underline{v}_p)\} = \{A\underline{v}_1, \dots, A\underline{v}_p\}.$$

But how can we "go from 2 to 1"? I.e.:

Given: $\{\underline{v}_1, \dots, \underline{v}_p\}$ and $\{T(\underline{v}_1), \dots, T(\underline{v}_p)\}$, Find: A .

Solution: we seek A s.t.

$$\underbrace{T(\underline{v}_1)}_{\underline{u}_1} = A\underline{v}_1, \dots, \underbrace{T(\underline{v}_p)}_{\underline{u}_p} = A\underline{v}_p, \Rightarrow$$

$$\underbrace{[\underline{u}_1, \dots, \underline{u}_p]}_{\underline{U}} = [A\underline{v}_1, \dots, A\underline{v}_p] = A \underbrace{[\underline{v}_1, \dots, \underline{v}_p]}_{\underline{V}}$$

$$\Rightarrow (U = AV) V^{-1} \Rightarrow UV^{-1} = A \cancel{(VV^{-1})} \rightarrow I$$

$$\Rightarrow \boxed{A = UV^{-1}}$$

Note 1: V^{-1} exists because $\{\underline{v}_1, \dots, \underline{v}_p\}$ are lin. indep., $\Rightarrow V$ is nonsingular.

Note 2: When $\{\underline{v}_1, \underline{v}_2\} = \{\underline{e}_1, \underline{e}_2\}$ (the natural basis (pp. 195-196 of book), then $V = I$. Then

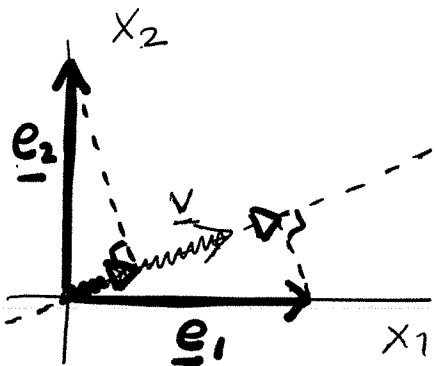
$$A = U \equiv [T(\underline{e}_1), T(\underline{e}_2)] \quad (\text{Thm. 15 in book})$$

Note 3: While $\{v_1, v_2\}$ are lin. indep.,

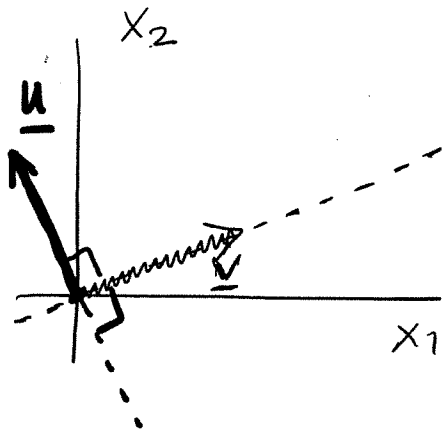
$\{u_1, u_2\} = A \cdot \{v_1, v_2\}$ may be lin. dep. This occurs when A is singular.

Ex. 7 The matrix of the projection (Ex. 4) is singular. \leftarrow Show this.

Sol'n: 1) First, it's clear that for any basis



(e.g., for $\{e_1, e_2\}$), its projections on the direction of v are parallel (since they are on the same line).



2) To demonstrate that the matrix A of the projection is singular, we need to show that for some $u \neq \underline{0}$,

$$A u \equiv P_v(u) = \underline{0}.$$

This is true for any $u \perp v$ (see figure).

One can write: $N(A) = \{x : x \perp v\}$.

Note: The matrices of l.t. in Ex. 1, 2 were nonsingular.

• **MUST READ** pp. 233-234 about null space and range of a l.t.

- range of T = all possible outcomes of T ;
- null space of T = all x s.t. $T(x) = \underline{0}$ (Ex. 7).

4) Composition of l. t.

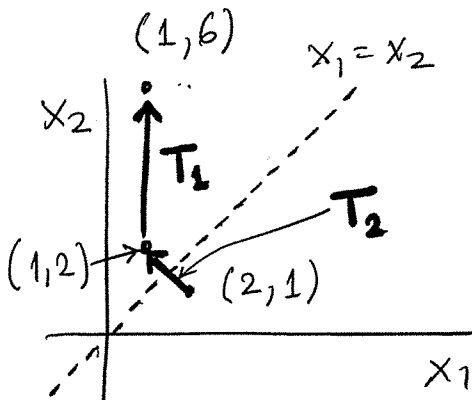
In sec. 1.5 we learned that $AB \neq BA$ for matrices in general. We can now illustrate this.

Ex. 8 Let T_1 be the vertical stretch by a factor of 3 in \mathbb{R}^2 , and T_2 be the reflection about $x_1 = x_2$. Find the composition of T_1 & T_2 and then of T_2 & T_1 , and verify that they yield different results.

Sol'n: 1) $T_2(\underline{x}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \underline{x}, \Rightarrow A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (P. 15-2)

$T_1\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = \begin{pmatrix} x_1 \\ 3x_2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \Rightarrow A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$.

2) $T_1 \circ T_2 \leftarrow$ composition of T_1 & T_2 , where T_2 is applied first:



$(T_1 \circ T_2)(\underline{x}) \equiv T_1(T_2(\underline{x}))$

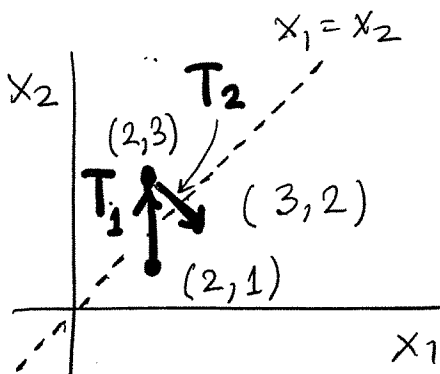
$T_2\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$T_1\left(T_2\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right]\right) = T_1\left[\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 6 \end{pmatrix}$

Via matrix multiplication:

$T_1 \circ T_2(\underline{x}) = A_1(A_2 \underline{x}) = \underline{A_1 A_2} \underline{x} = \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix} \underline{x}$

3) $T_2 \circ T_1 \leftarrow$ composition of T_2 & T_1 , where T_1 is applied first:



$T_1\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

$T_2 \circ T_1\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) = T_2\left[\begin{pmatrix} 2 \\ 3 \end{pmatrix}\right] = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

Via matrix multiplication:

$T_2 \circ T_1(\underline{x}) = \underline{A_2 A_1} \underline{x} = \begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix} \underline{x}$

This example explains visually why
 $A_1 A_2 \neq A_2 A_1$.

In words, the result of a composition of two l.t. depends on the order in which the l.t. are performed.

Note: This visual example also allows us to see when two transformations (and hence their matrices) will commute. For example, if we perform two consecutive rotations, by angles θ_1 and θ_2 , about the origin, the result will not depend in which order we perform these rotations. So, if R_θ is the matrix of rotation by angle θ (see p. 236), then $R_{\theta_1} \cdot R_{\theta_2} = R_{\theta_2} \cdot R_{\theta_1}$.

⑤ Orthogonal l.t.

Def: A l.t. $T(x)$ s.t.

$$\|T(x)\| = \|x\| \text{ for all } x$$

is called an orthogonal l.t.

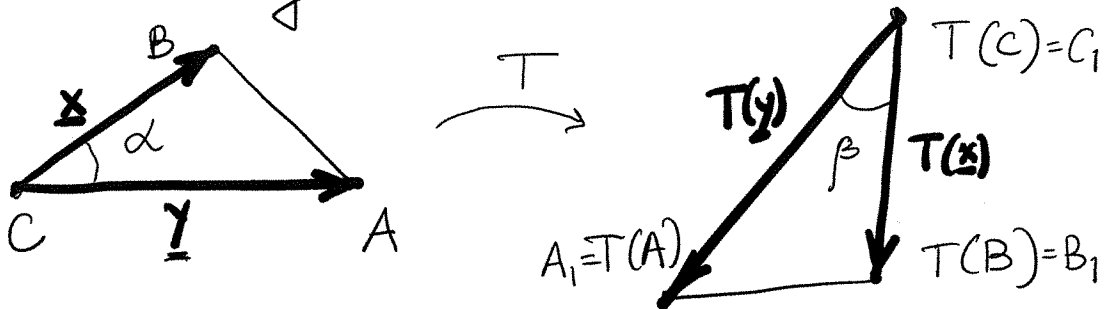
Ex. 9 A reflection (about any line through the origin) and a rotation about the origin are orthogonal l.t. See pp. 236-237 in book. Moreover, they are the only orthogonal l.t. in \mathbb{R}^2 . In \mathbb{R}^3 , orthogonal l.t. also include only rotations and reflections, but there are more possibilities there.

Claim: An orthogonal l.t. preserves the angle between any two vectors. I.e.,

$$\left(T = \begin{array}{l} \text{orthogonal l.t.} \end{array} \right) \Rightarrow \left(\angle(T(x), T(y)) = \angle(x, y) \right)$$

for any x, y .

- One can prove this using the concept of the dot product, similar to the proof of Thm. 16. However, there is a much easier visual proof: Consider any $\triangle ABC$.



T transforms it into $\triangle T(A)T(B)T(C) \equiv \triangle A_1B_1C_1$. Since T is orthogonal, the sides of $\triangle A_1B_1C_1$ have the same lengths as the respective sides of $\triangle ABC$. Then by comparing the congruent $\triangle A_1B_1C_1 \cong \triangle ABC$, we have $\angle A_1C_1B_1 = \angle ACB$ (i.e., $\angle(T(x), T(y)) = \angle(x, y)$), and similarly for the other angles. q.e.d.

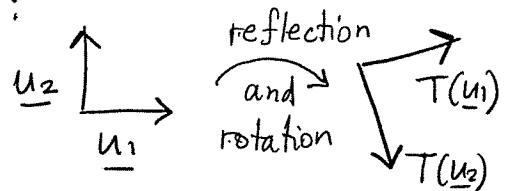
Corollary (= Thm. 16): An orthogonal l.t. transforms an orthonormal basis into another orthonormal basis:

$$\left(T = \begin{array}{l} \text{orthogonal l.t.} \end{array} \right) \text{ and } \left(\begin{array}{l} \{ \underline{u}_1, \underline{u}_2 \} = \text{orthonormal,} \\ \text{i.e. } \| \underline{u}_1 \| = \| \underline{u}_2 \| = 1, \\ \underline{u}_1 \perp \underline{u}_2 \end{array} \right)$$

$$\Downarrow$$

$$\left(\begin{array}{l} \{ T(\underline{u}_1), T(\underline{u}_2) \} = \text{orthonormal, i.e.} \\ \| T(\underline{u}_1) \| = \| T(\underline{u}_2) \| = 1 \text{ and } T(\underline{u}_1) \perp T(\underline{u}_2) \end{array} \right)$$

Note 1: You can easily visualize this in \mathbb{R}^2 , where reflections and rotations (which are the only orthogonal l.t.; see p. 15-10) preserve the orthonormality of a basis:



Note 2: The same

claim also holds in any \mathbb{R}^n .

Note 3 The matrix of an orthogonal l.t. is called the orthogonal matrix. We will consider them (defined differently, but called the same name) in Sec. 4.7.

One can also show, using a manipulation of orthogonal matrices, that a composition of two orthogonal l.t. is again an orthogonal l.t. (the the figure above on this page).