Example 3.48. Given the matrix

$$C = \begin{bmatrix} 1 & -1 & 2 & -2 \\ 2 & 0 & 1 & 0 \\ 5 & -3 & 7 & -6 \\ 1 & 1 & -1 & 3 \end{bmatrix},$$

find a basis and the dimension of the

(a) column space of C

(b) row space of C

(c) solution space of C.

3.8 Change of basis

We continue our investigation of the benefits of having bases for vector spaces and look into the effects of changing basis.

Recall that an ordered basis $\mathcal{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ of a vector space V gives rise to an isomorphism

$$\varphi_{\mathcal{B}} \colon V \to \mathbb{R}^n$$

defined by taking coordinates with respect to $\mathcal{B}$: $\varphi_{\mathcal{B}}(\mathbf{w}) = [\mathbf{w}]_{\mathcal{B}}$.

3.8.1 Effect of change of basis on coordinates

Suppose we are given a second ordered basis $\mathcal{C} = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ for V.

This gives rise to another isomorphism $\varphi_{\mathcal{C}}$ from V to $\mathbb{R}^n$, which we can fit into a diagram

So we end up with a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$, which we know corresponds to an $n \times n$ matrix.

We denote this matrix $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and call it *the change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$*.

The change of basis matrix is straightforward to compute:

$$P_{\mathcal{C}\leftarrow\mathcal{B}} =$$

Example 3.49. In $V = \mathbb{R}^2$, write down the change of basis matrix from $\mathcal{B}$ to $\mathcal{S}$, where $\mathcal{B} = ((1,1),(1,-1))$ and $\mathcal{S} = ((1,0),(0,1))$, and use it to compute $[\mathbf{v}]_{\mathcal{S}}$, given that $[\mathbf{v}]_{\mathcal{B}} = \cancel{(1,1)}.$ $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$P_{S\leftarrow B} = \begin{bmatrix} [(1,1)]_S & [(1,-1)]_S \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

在S为基情况下 v1, v2的坐标不变

$$[v]_S = P_{C\leftarrow B} \cdot [v]_B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

How do you think the two change of basis matrices $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and $P_{\mathcal{B}\leftarrow\mathcal{C}}$ are related?

3.8.2 Matrix representation of a linear transformation

Now suppose we have a vector space V with ordered basis $\mathcal{B}$, a vector space W with ordered basis $\mathcal{C}$, and a linear transformation $T\colon V \to W$.

This gives rise to a diagram

So we end up with a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$, which we know corresponds to an $m \times n$ matrix.

We denote this matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ and call it *the matrix of T with respect to the ordered bases $\mathcal{B}$ and $\mathcal{C}$*.

It has the property that

$$[T(\mathbf{v})]_{\mathcal{C}} = [T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} \qquad \text{for all } \mathbf{v} \in V.$$

In the special case where $W = V$ and $\mathcal{C} = \mathcal{B}$, we write simply $[T]_{\mathcal{B}}$ instead of $[T]_{\mathcal{B}\leftarrow\mathcal{B}}$.

It has the property that

$$[T(\mathbf{v})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} \qquad \text{for all } \mathbf{v} \in V.$$

The matrix representation of T is straightforward to compute:

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}} =$$

Example 3.50. Consider the linear transformation $T\colon \mathcal{P}_1 \to \mathcal{P}_2$, $T(f) = (x+2)f$. Find the matrix of T with respect to the ordered bases $(1, x)$ and $(1, x, x^2)$. Determine the image of $2 + 3x$ under T in two ways.

3.8.3 Effect of change of basis on matrix representations

We will work out in detail the special case of $T\colon V \to V$, which is prominent in applications. The general case $T\colon V \to W$ can be treated using the same approach.

Suppose V is a vector space and $T\colon V \to V$ is a linear transformation. Given an ordered basis $\mathcal{B}$, we get a matrix representation $[T]_{\mathcal{B}}$. Given another ordered basis $\mathcal{C}$, we get another matrix representation $[T]_{\mathcal{C}}$.

Can we relate the two matrices?

We conclude that

$$[T]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}\,[T]_{\mathcal{B}}\,P_{\mathcal{B}\leftarrow\mathcal{C}}.$$

Example 3.51. Consider the linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by the matrix

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

and the ordered bases $\mathcal{S}$ (the standard basis) and $\mathcal{C} = ((1,1,1),(1,1,0),(1,0,0))$.

$T(v) = Av$

$\downarrow$

$[T]_S$

$P_{S\leftarrow C} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

$\downarrow$

Take C under the standard basis

$\to$ put C in column

check

$[T]_C = \begin{bmatrix} \vdots & \vdots & \vdots \\ [T(v_1)]_C & [T(v_2)]_C & [T(v_3)] \\ \vdots & \vdots & \vdots \end{bmatrix}$

$P_{C\leftarrow S} = P_{S\leftarrow C}^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$

$T(w_1) = Aw_1 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = a_1\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + a_2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + a_3\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

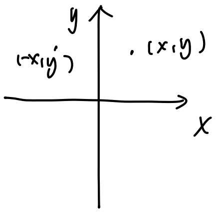
$\to \boxed{a_1 = 4 \quad a_2 = -1 \quad a_3 = -1}$

so $[T]_C = P_{C\leftarrow S}\,[T]_S\,P_{S\leftarrow C}$.

3.9 Linear transformations and geometry of $\mathbb{R}^2$

Many of the geometric transformations on $\mathbb{R}^2$ are linear.

Example 3.52 (Reflections)**.** Consider $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by reflection in the y-axis.



$T(x,y) = (-x, y)$ for all $(x,y) \in R^2$

check it's linear

(a) $T((x_1,y_1)+(x_2,y_2)) = T((x_1+x_2, y_1+y_2)) = -(-(x_1+x_2), y_1+y_2)$

$T(x_1,y_1) + T(x_2,y_2) = (-x_1, y_1) + (-x_2, y_2)$

$= (-x_1-x_2, y_1+y_2)$

(b) $= T(\lambda(x,y)) = T(\lambda x, \lambda y) = (-\lambda x, \lambda y)$

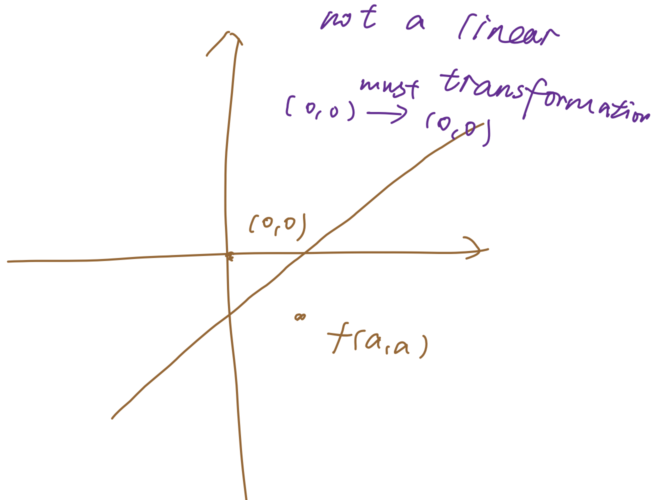
$\lambda T(x,y) = \lambda(-x,y) = (-\lambda x, \lambda y)$

More generally, reflection in any line passing through the origin is a linear transformation.

Matrix rep of $T: R^2 \to R^2$ with respect to standard basis of $\mathcal{B}$ (e_1, e_2) of R^2

$$[T]_{\mathcal{B}} = \left[T(e_1) \;,\; T(e_2) \right] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$K(T) = \{(0,0)\}$$



Example 3.53 (Dilations, contractions)**.** Consider $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by scaling everything by a factor of 3.

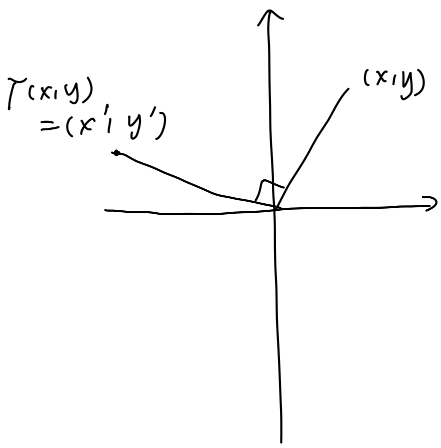
$T(x, y) = (3x, 3y)$

linear transformation

$[T]_{\gamma} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$

$\ker(T) = \{(0,0)\}$

Example 3.54 (Rotations)**.** Consider $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by anti-clockwise rotation around the origin by an angle of $\frac{\pi}{2}$.



This is a linear transformation

$[T]_{\gamma} = \begin{bmatrix} \vdots & \vdots \\ T(e_1) & T(e_2) \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

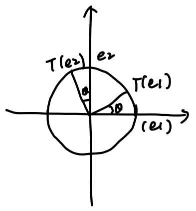
$T(x, y) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

$= \begin{bmatrix} -y \\ x \end{bmatrix}$

More generally, rotation around the origin by any angle θ is a linear transformation.

It is useful to have a matrix representation for this:

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

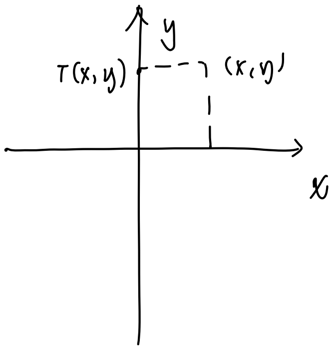


$$[T]_J = \begin{bmatrix} T(e_1) & T(e_2) \end{bmatrix} = \begin{bmatrix} \cos\theta & \cos(\theta+\frac{\pi}{2}) \\ \sin\theta & \sin(\theta+\frac{\pi}{2}) \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$\det[T]_J = 1 \rightarrow$ linear transformation

Example 3.55 (Orthogonal projections). Consider $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by orthogonal projection onto the y-axis.



$T(x, y) = (0, y)$

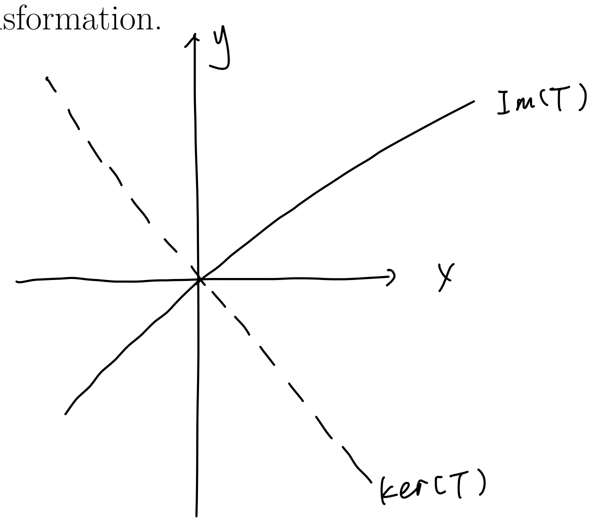
linear transformation

$$[T]_J = \begin{bmatrix} T(e_1) & T(e_2) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$\ker(T) = \{(x, 0) \mid x \in \mathbb{R}\}$ dim 1

$\operatorname{Im}(T) = \{(0, y) \mid y \in \mathbb{R}\}$ dim 1

More generally, orthogonal projection onto any line passing through the origin is a linear transformation.



Example 3.56 (Translations)**.** Consider $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by translation by the vector $(1, -2)$.

$$S(x,y) = (x,y) + (1,-2) = (x+1, y-2)$$

Note $S(0,0) = (1,-2)$ so S is not linear transformation

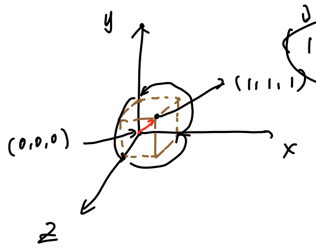
? not closed on scalar multiplication

Translations are basically never linear transformations (the only one that is linear is translation by $(0,0)$).

More generally, any $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that $S(0,0) \neq (0,0)$ is not linear.

All the types of linear transformations we discussed have generalisations to $\mathbb{R}^3$ (and, with appropriate care, higher $\mathbb{R}^n$). Finding matrix representatives follows the same general principles but can be challenging to implement.

Example 3.57. Consider the cube of side length one with a vertex at $(0,0,0)$ and another at $(1,1,1)$. Let L denote the diagonal joining these two vertices, and let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ be the rotation about L by $\frac{2\pi}{3}$.



1圈 $\times \frac{1}{3}$

$$[T]_{\mathcal{E}} = \begin{bmatrix} T(e_1) & T(e_2) & T(e_3) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

3.10 Linear algebra over $\mathbb{F}_2$ and coding theory

A *binary linear code* is a subspace C of the vector space $\mathbb{F}_2^n$.

The vectors in C are called its *codewords*.

A matrix A with entries in $\mathbb{F}_2$ is called a *check matrix* for C if $\ker(A) = C$.

Example 3.58.

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

is a check matrix for

$$C = \{(0,0,0,0), (1,0,1,0), (0,1,0,1), (1,1,1,1)\}.$$

The name "check matrix" indicates what it is used for: If we receive a word, say $\mathbf{v} = (1,1,0,1)$, we can check whether it is a valid codeword by verifying $\boxed{A\mathbf{v} = \mathbf{0}}$.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ ? \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ so $(1,1,0,1)$ is not a codeword

The vector space $\mathbb{F}_2^n$ has the theoretical capacity of distinguishing 2^n different pieces of information. Why would we want to restrict to a subspace C, which will clearly reduce this capacity?

ASCII code

So, in the real world, we want our codes to detect transmission errors.

Our code C from Example 3.58 has this property, in that it can detect one bit flip in any group of four bits.

What do we do when we detect an error? One possibility is to ask for the message to be transmitted again and hope that it makes it through intact this time.

Or we can design our codes to correct errors: When we receive a word, instead of discarding it if it is not a codeword, we can replace it by the most likely-looking codeword.

To make this precise, we need a good notion of distance in this setting.

The *Hamming distance* between $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{F}_2^n$ and $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{F}_2^n$ is the number $d(\mathbf{u}, \mathbf{v})$ of indices i for which $u_i \neq v_i$.

if $u_1 = v_1$ ✓
$u_2 \neq v_2$ ✗
⋮
$u_n = v_n$ ✓
⇒ total number of "$u_i \neq v_i$"

Example 3.59. In $\mathbb{F}_2^4$:

$d((0,1,1,0),(1,0,1,0)) =$ $1+1=2$

$d((0,0,0,0),(1,1,1,1)) =$ 4

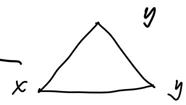
Proposition 3.60. The Hamming distance $d\colon \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{Z}$ satisfies the following properties:

(a) $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$, and $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$

(b) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$

(c) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}_2^n$.

triangle inequality in $\mathbb{R}^2$



proof:

for any $n \in \mathbb{N}_0$, $d: \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{Z}_{\geq 0}$ satisfie (c)

By induction on n

Base case $n=1$ $\quad x, y, z \in \mathbb{F}_2$

x	y	z	$d(x,y)$	$d(x,z)+d(z,y)$
0	0	0	0	$0+0=0$
0	0	1	0	$1+1=2$
0	1	0	1	$0+1=1$
0	1	1	1	⋮
1	0	0	1	
1	0	1	1	
1	1	0	0	
1	1	1	0	

$x = (x_0, \boxed{x_1, \cdots x_n})$ $\quad x'$

$y = (y_0, \boxed{y_1, \cdots y_n})$ $\quad y'$

$z = (z_0, \boxed{z_1, \cdots z_n})$. $\quad z'$ $\quad x_i, y_i, z_i \in \mathbb{F}_2$

Induction step.

Fix $n \in \mathbb{N}$, assume true for n

Let $x, y, z \in \mathbb{F}_2^{n+1}$

$d(x,y) = d(x_0, y_0) + d(x', y')$

$d(x,z) = d(x_0, z_0) + d(x', z')$

$d(y,z) = \underbrace{d(y_0, z_0)}_{\text{base case}} + \underbrace{d(y', z')}_{\text{induction hyp}}$

$d(x,y) \leq d(x,z) + d(y,z)$ for $n+1$

From now on, when we receive a word $\mathbf{x}$, we will look for the codeword that is closest to $\mathbf{x}$ and take that to have been the original message.

Example 3.61. Consider the code C with check matrix

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Then

$$C = \{(0,0,0,0,0,0), (1,0,1,0,0,1), (1,1,0,1,0,0), (1,1,1,0,1,1),$$
$$(1,0,0,1,1,0), (0,1,0,0,1,0), (0,1,1,1,0,1), (0,0,1,1,1,1)\}$$

Suppose we receive $\mathbf{v} = (0,0,1,0,0,0)$, then assume that the original message is $(0,0,0,0,0,0)$

What if we receive $\mathbf{w} = (1,0,1,0,1,1)$? Then $d(w, 101001)$

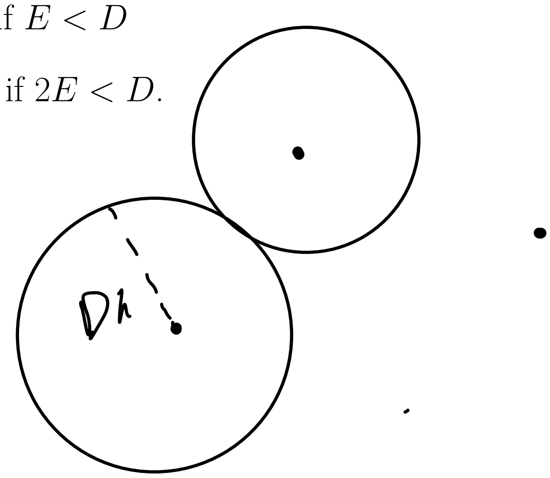
$= d(w, 111011) = 1$

Guess: In a good code, codewords are as far apart as possible.

Theorem 3.62. Let C be a code with minimum distance D between any two distinct codewords. Then

(a) C can detect E errors if $E < D$

(b) C can correct E errors if $2E < D$.



Example 3.63. The *Hamming* $(7,4)$*-code* is defined by the check matrix

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The codewords are

0000000	1101001	0101010	1000011
1001100	0100101	1100110	0001111
1110000	0011001	1011010	0110011
0111100	1010101	0010110	1111111

$\frac{16\times 15}{2}$ pairs

The minimum distance for this code is 3

the minimum of all pair of codeword

↓

So the code can

detect 2 errors

correct 1 error | the minimum of any codeword to the origin ⇒ just 15 codeword (compare to 0000000)

test