a) ~~$t_i w_{ij}$~~ $\underset{5\times 7}{X} \approx \underset{5\times1}{T}\,\underset{1\times7}{W} \to \underset{1\times 7}{\underline{w_1}^T}$

$$\underset{5\times1}{\underline{t_1}} = \frac{1}{\sqrt5}\begin{bmatrix}1\\\vdots\\1\end{bmatrix} \cdot [w_{11}\ w_{12} \cdots w_{17}]$$

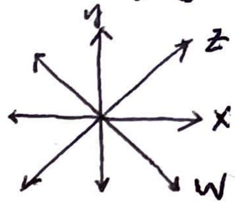
$$= \begin{bmatrix} x_{:,1} & x_{:,2} & \cdots & x_{:,7} \\ \vdots & \vdots & & \vdots \\ x_{:,1} & x_{:,2} & \cdots & x_{:,7}\end{bmatrix}$$

avg. of col X

$\Rightarrow w_{1j} = \sqrt5\, x_{:,j}$

Activity 5 △ Item 1 $z = \begin{bmatrix}1\\1\end{bmatrix}, w = \begin{bmatrix}1\\-1\end{bmatrix}$

c) span $\{z, w\}$ in $\mathbb{R}^2$



(on x-y plane)

d) $z^T w = [1, 1]\begin{bmatrix}1\\-1\end{bmatrix} = 0$, so z, w are orthogonal.

e) $|z| = \sqrt2 = |w| \neq 1 \to z, w$ are not orthonormal

$\to \left|\frac{1}{\sqrt2}(z)\right| = \left|\frac{1}{\sqrt2}(w)\right| = 1$

$\Rightarrow z'' = \begin{bmatrix}\frac{1}{\sqrt2}\\\frac{1}{\sqrt2}\end{bmatrix}, w' = \begin{bmatrix}\frac{1}{\sqrt2}\\-\frac{1}{\sqrt2}\end{bmatrix}$

△ Item 2 a) $x_2 = x_1 + 1$

$(0, 1), (-1, 0)$

b) the line does not go through origin.

△ Item 3

a) $\underset{5\times7}{X} \approx \underset{5\times k}{\underline{T}}\,\underset{k\times7}{\underline{W}}$

rank$(X) = 5$

b) rank-r taste matrix with orthonormal columns

$\to$ No. of columns $= r$

$\underline{T} : 5\times r$, $\underline{W} : r \times 7$

In [5]:
```
import numpy as np
from scipy.io import loadmat
import matplotlib.pyplot as plt
# import statistics as st

in_data = loadmat('movie.mat')
#loadmat() loads a matlab workspace into a python dictionary, where the names of the var
#in the dictionary.  To see what variables are loaded, uncomment the line below:
print([key for key in in_data])

M = in_data['M']
print(M)
```

```
['__header__', '__version__', '__globals__', 'M', 'ans']
[[ 4  7  2  8  7  4  2]
 [ 9  3  5  6 10  5  5]
 [ 4  8  3  7  6  4  1]
 [ 9  2  6  5  9  5  4]
 [ 4  9  2  8  7  4  1]]
```

In [3]:
```
np.linalg.matrix_rank(M)
```

Out[3]:
```
5
```

In [19]:
```
# M_T = M.T
# print(M_T)
avg_M = [list(M.mean(0))]*5
print(avg_M)
avg_ls = list(M.mean(0))
avg_ls
```

```
[[6.0, 5.8, 3.6, 6.8, 7.8, 4.4, 2.6], [6.0, 5.8, 3.6, 6.8, 7.8, 4.4, 2.6], [6.0, 5.8, 3.
6, 6.8, 7.8, 4.4, 2.6], [6.0, 5.8, 3.6, 6.8, 7.8, 4.4, 2.6], [6.0, 5.8, 3.6, 6.8, 7.8,
4.4, 2.6]]
```

Out[19]:
```
[6.0, 5.8, 3.6, 6.8, 7.8, 4.4, 2.6]
```

In [18]:
```
cons = (5)**(0.5)
w1_T = [x * cons for x in avg_ls]
w1_T
```

Out[18]:
```
[13.416407864998739,
 12.96919426949878,
 8.049844718999244,
 15.20526224699857,
 17.44133022449836,
 9.838699100999076,
 5.813776741499454]
```

In [20]:
```
np.subtract(M, avg_M)
```

Out[20]:
```
array([[-2. ,  1.2, -1.6,  1.2, -0.8, -0.4, -0.6],
       [ 3. , -2.8,  1.4, -0.8,  2.2,  0.6,  2.4],
       [-2. ,  2.2, -0.6,  0.2, -1.8, -0.4, -1.6],
       [ 3. , -3.8,  2.4, -1.8,  1.2,  0.6,  1.4],
       [-2. ,  3.2, -1.6,  1.2, -0.8, -0.4, -1.6]])
```

f) Do you see any patterns in the residual? Briefly describe them qualitatively.

Yes, there is a noticeable pattern in the residuals. It appears that individuals who favor sci-fi movies tend to have positive residuals for sci-fi movies (1st, 3rd, and 5th), while showing negative residuals for romance movies, and vice versa.

In [ ]: