EE2211 Introduction to Machine Learning

T14 & T22, Chua Dingjuan elechuad@nus.edu.sg
Slides @ tiny.cc/ee2211tut

Tutorial 4

Systems of Linear Equations

Brief Summary of Key Points

Matrix Multiplication (Dot Pdt)

$$\mathbf{X}\mathbf{w} = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,1} & x_{3,2} & x_{3,3} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$= \begin{bmatrix} x_{1,1}w_1 + x_{1,2}w_2 + x_{1,3}w_3 \\ x_{2,1}w_1 + x_{2,2}w_2 + x_{2,3}w_3 \\ x_{3,1}w_1 + x_{3,2}w_2 + x_{3,3}w_3 \end{bmatrix}$$

Size Rules

$$\underset{\mathbf{A}}{\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 2 & 5 \end{bmatrix}} \cdot \underset{\mathbf{B}}{\begin{bmatrix} 1 & 3 & 2 & 2 \\ 2 & 4 & 5 & 1 \end{bmatrix}}$$

($\#row_A$ x $\#col_A$) * ($\#row_B$ x $\#col_B$) = ($\#row_A$ x $\#col_B$)

$$(m \times n) \cdot (n \times k) = (m \times k)$$

product is defined

Matrix Inverse & Det

(X is invertible if all rows or columns of X are linearly independent)

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \operatorname{adj}(\mathbf{A})$$

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{ then } \operatorname{adj}(\mathbf{A}) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\det(\mathbf{A}) = |\mathbf{A}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

If det(A) = 0 ➔ not invertible (singular)

(square)

Brief Summary of Key Points

- **System of Linear Equations**

$$\mathbf{X}\mathbf{w} = \mathbf{y}$$

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,d} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_d \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

X is of *m-by-d* dimension
m rows by d cols

d rows by 1 col

m rows by 1 col

Even Determined m = d	Over Determined m > d	Under Determined m < d
Exact solution	No exact solution = approximate solution	infinite number of solutions = constrained solution $\mathbf{w} = \mathbf{X}^T\mathbf{a}$
$\mathbf{w} = \mathbf{X}^{-1}\mathbf{y}$	**Left inv** $\mathbf{X}^{\dagger} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ $\hat{\mathbf{w}} = \mathbf{X}^{\dagger}\mathbf{y}$ Least squares	**Right inv** $\mathbf{X}^{\dagger} = \mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^{-1}$ $\hat{\mathbf{w}} = \mathbf{X}^{\dagger}\mathbf{y}$ Least norms

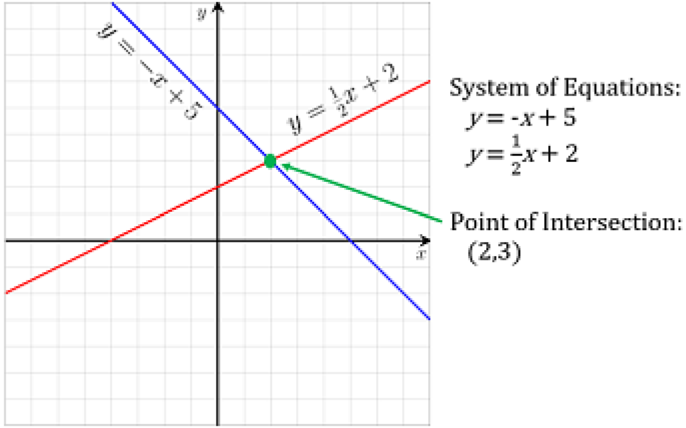
If both XX^T, X^TX are not invertible, no solution (test for invertible or not using det)

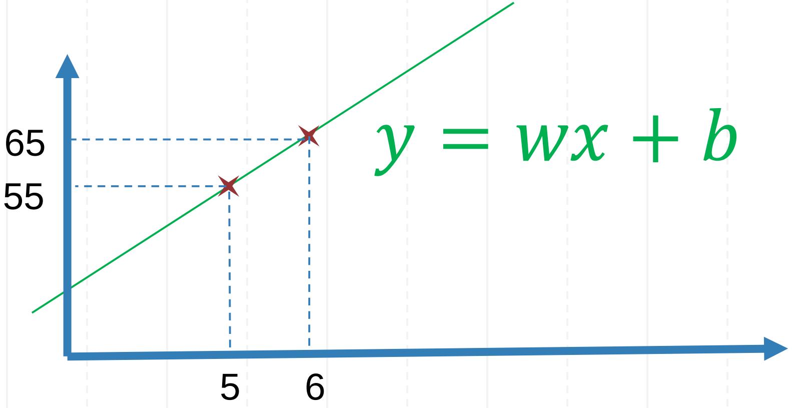
Brief Summary of Key Points

- **Even Determined**

$$\mathbf{Xw} = \mathbf{y}$$

Even Determined m = d
Exact solution
$\mathbf{w} = \mathbf{X}^{-1}\mathbf{y}$





numpy - arrays

- `numpy` : "Numeric(al) Python" - supports large, multi-dimensional arrays and matrices & high-level mathematical functions to operate on these arrays
- `numpy arrays` : multidimensional array object that is a powerful data structure for efficient computation of arrays and matrices
- `np.array( [ [row 1] , [row 2] ] )`
  - Creating a numpy array

```
>>> np.array( [ [1,2] , [5,7] ])
array([[1, 2],
       [5, 7]])
```

- `np.ones ((#rows, #cols)), np.zeros np.full ((#rows, #cols),value)`
  - Creating default arrays

```
>>> np.zeros((2,3))
array([[0., 0., 0.],
       [0., 0., 0.]])
```

```
>>> np.ones((2,3))
array([[1., 1., 1.],
       [1., 1., 1.]])
```

```
>>> np.full((2,3),2)
array([[2, 2, 2],
       [2, 2, 2]])
```

- `np.eye(#row), np.identity(#row)`
  - Creating identity matrices

```
>>> np.identity(2)
array([[1., 0.],
       [0., 1.]])
```

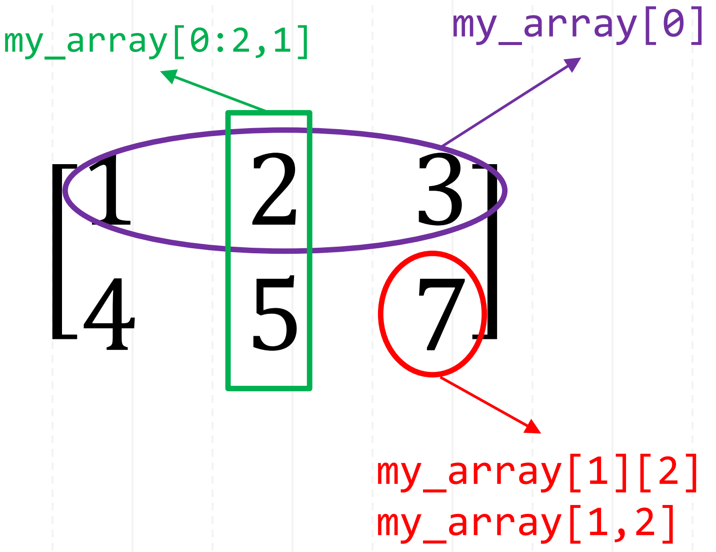
https://www.datacamp.com/community/tutorials/python-numpy-tutorial

numpy - indexing

- `my_array = np.array( [ [1,2,3] , [4,5,7] ] )`

Note:zero-indexing…

- `my_array[0]`
  - Zeroth row
- `my_array[1][2]`
  - Element at row 1 column 2
- `my_array[1,2]`
  - Element at row 1 column 2
- `my_array[0:2,1]`
  - items at row 0 and 1, column 1




```
a[start:end] # items start through the end (but the end is not included!)
a[start:]    # items start through the rest of the array
a[:end]      # items from the beginning through the end (but the end is not included!)
```

https://www.datacamp.com/community/tutorials/python-numpy-tutorial

numpy – mathematical functions

- `np.dot(A,B)`
- `A.dot(B)`
  - Dot Product
- `np.add(), np.subtract()`
  - Mathematical operations for matrices
- `np.transpose()   or   ().T`
- `np.linalg.inv()`
  - Calculating multiplicative inverse of matrix
- `np.linalg.det()`
  - Calculating determinant of matrix

NOTE
`np.multiply() ,  *`
is NOT the usual matrix dot product.

A

A(0,0)	A(0,1)
A(1,0)	A(1,1)

B

B(0,0)	B(0,1)
B(1,0)	B(1,1)

A(0,0) * B(0,0)	A(0,1) * B(0,1)
A(1,0) * B(1,0)	A(1,1) * B(1,1)

$\boxed{\textbf{numpy.multiply(A, B)}}$

https://numpy.org/doc/stable/reference/generated/numpy.dot.html#numpy.dot

Dot Products

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

```
>>> A
array([[1, 2],
       [3, 4]])
>>> B
array([[2, 2],
       [2, 2]])
```

$$AB = A.B$$
$$= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} 6 & 6 \\ 14 & 14 \end{bmatrix}$$

- np.dot(A,B)
- A @ B
- np.matmul(A,B)

```
>>> np.dot(A,B)
array([[ 6,  6],
       [14, 14]])
>>> A.dot(B)
array([[ 6,  6],
       [14, 14]])
>>> A@B
array([[ 6,  6],
       [14, 14]])
>>> np.matmul(A,B)
array([[ 6,  6],
       [14, 14]])
```

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \; B = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

$$\text{Multiply} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$$

- np.multiply(A,B)

  or

- A * B

```
>>> np.multiply(A,B)
array([[2, 4],
       [6, 8]])
>>> A * B
array([[2, 4],
       [6, 8]])
```

https://numpy.org/doc/stable/reference/generated/numpy.dot.html#numpy.dot

Discussion of Solutions 4

Q1-4,5,6,7,8

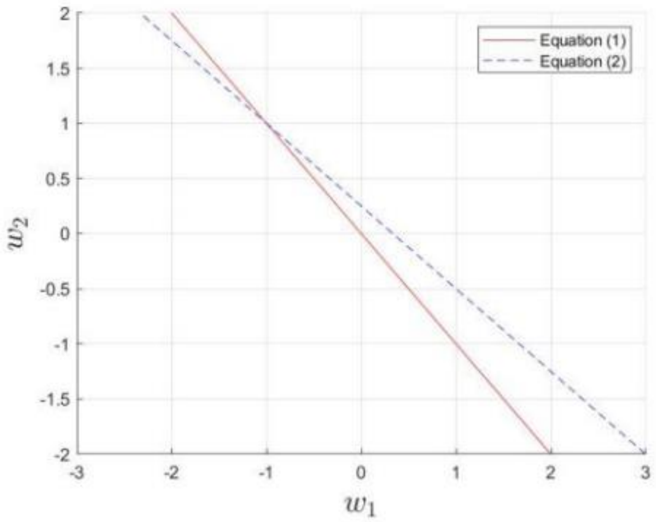
Question1

Given $\mathbf{Xw} = \mathbf{y}$ where $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

(a) What kind of system is this? (even-, over- or under-determined?)

(b) Is $\mathbf{X}$ invertible? Why?

(c) Solve for $\mathbf{w}$ if it is solvable.



Answer:

(a) This is an even-determined system.

(b) $\det(\mathbf{X}) = 1 \times 4 - 1 \times 3 = 1 \neq 0$. $\mathbf{X}^{-1} = \frac{1}{1}\begin{bmatrix} 4 & -1 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ -3 & 1 \end{bmatrix}$.

(c) $\hat{\mathbf{w}} = \mathbf{X}^{-1}\mathbf{y} = \begin{bmatrix} 4 & -1 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

```python
import numpy as np

m_list = [[1, 1], [3, 4]]
X = np.array(m_list)

inv_X = np.linalg.inv(X)
y = np.array([[0], [1]])

deter = np.linalg.det(X)
print(deter)

w = inv_X.dot(y)

w = np.dot(inv_X, y)
```

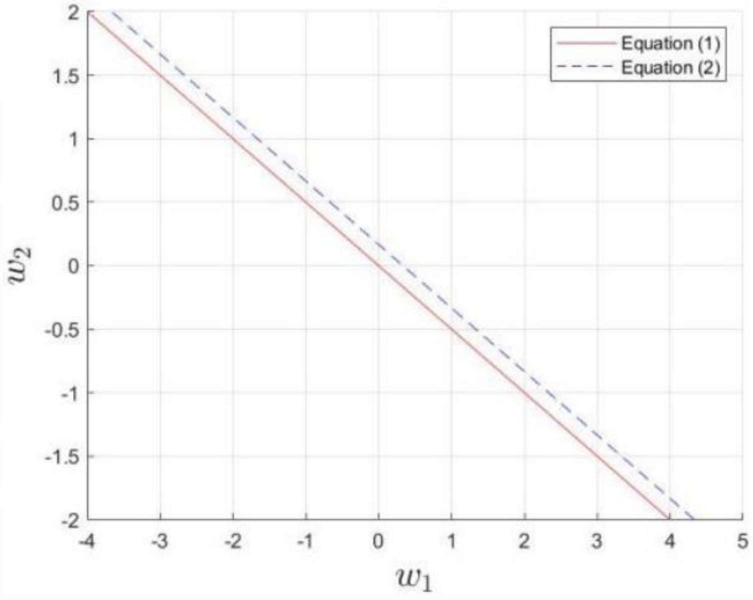
Question2

Given $\mathbf{Xw} = \mathbf{y}$ where $\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

(a) What kind of system is this? (even-, over- or under-determined?)

(b) Is $\mathbf{X}$ invertible? Why?

(c) Solve for $\mathbf{w}$ if it is solvable.



Answer:

(a) This is an even-determined system.

(b) $\mathbf{X}$ is NOT invertible since the determinant of $\mathbf{X}=1\times6 - 2\times3 = 0$.

(c) There is no solution for $\mathbf{w}$ since the rows/columns of $\mathbf{X}$ are inter-dependent. The two lines shown in the plot are parallel and has no intersection.

Question3

Given $\mathbf{X}\mathbf{w} = \mathbf{y}$ where $\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 1 & -1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 0 \\ 0.1 \\ 1 \end{bmatrix}$.

(a) What kind of system is this? (even-, over- or under-determined?)

(b) Is $\mathbf{X}$ invertible? Why?

(c) Solve for $\mathbf{w}$ if it is solvable.

```
import numpy as np
m_list = [[1, 2], [2, 4], [1, -1]]
X = np.array(m_list)

inv_XTX = np.linalg.inv(X.transpose().dot(X))

pinv = inv_XTX.dot(X.transpose())

y = np.array([[0], [0.1], [1]])

w = pinv.dot(y)
print(w)
```

w = inv(X.T @ X) @ X.T @ y

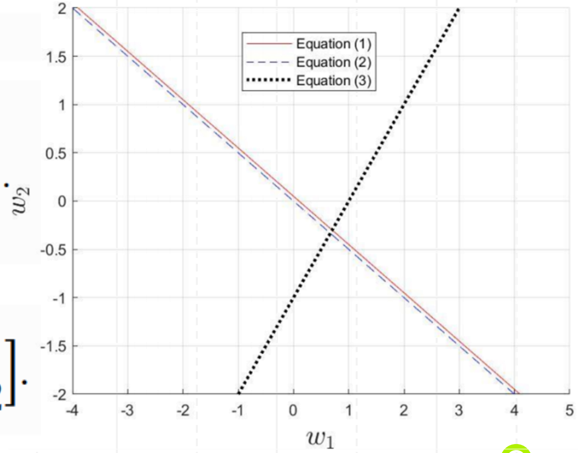
Answer:

(a) This is an over-determined system.

(b) $\mathbf{X}$ is NOT invertible but $\mathbf{X}^T\mathbf{X}$ is. The determinant of $\mathbf{X}^T\mathbf{X}$ $= 6\times 21 - 9\times 9 = 45$.

(c) An approximated solution is given by

$$\hat{\mathbf{w}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} = \begin{bmatrix} 0.4667 & -0.2 \\ -0.2 & 0.1333 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0.1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.68 \\ -0.32 \end{bmatrix}.$$



Question4

Given $\mathbf{X}\mathbf{w} = \mathbf{y}$ where $\mathbf{X} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

(a) What kind of system is this? (even-, over- or under-determined?)

(b) Is $\mathbf{X}$ invertible? Why?

(c) Solve for $\mathbf{w}$ if it is solvable.

```
import numpy as np
from numpy.linalg import inv

X = np.array([[1, 0,1,0], [1,-1,1,-1], [1, 1,0,0]])
y = np.array([[1],[0],[1]])

det = np.linalg.det(X.dot(X.transpose()))

inv_XTX = np.linalg.inv(X.dot(X.transpose()))

pinv = X.transpose().dot(inv_XTX)

w=pinv.dot(y)
```

Answer:

(a) This is an under-determined system.

(b) $\mathbf{X}$ is NOT invertible but $\mathbf{X}\mathbf{X}^T$ is.

The determinant of $\mathbf{X}\mathbf{X}^T = \det\left(\begin{bmatrix} 2 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 2 \end{bmatrix}\right) = 2\begin{vmatrix} 4 & 0 \\ 0 & 2 \end{vmatrix} - 2\begin{vmatrix} 2 & 0 \\ 1 & 2 \end{vmatrix} + 1\begin{vmatrix} 2 & 4 \\ 1 & 0 \end{vmatrix} = 2\times 8 - 2\times 4 + (-4) = 4$.

(c) $\hat{\mathbf{w}} = \mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 0.75 & 0.5 \\ -1 & 0.5 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$

Question5

Given $\mathbf{w}^T\mathbf{X} = \mathbf{y}^T$ where $\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

(a) What kind of system is this? (even-, over- or under-determined?)

(b) Is $\mathbf{X}$ invertible? Why?

(c) Solve for $\mathbf{w}$ if it is solvable.

Answer:

(a) This is an even-determined system.

(b) $\mathbf{X}$ is NOT invertible since the determinant of $\mathbf{X} = 1 \times 6 - 2 \times 3 = 0$.

(c) There is no solution for $\mathbf{w}$ (two parallel lines).

Question6

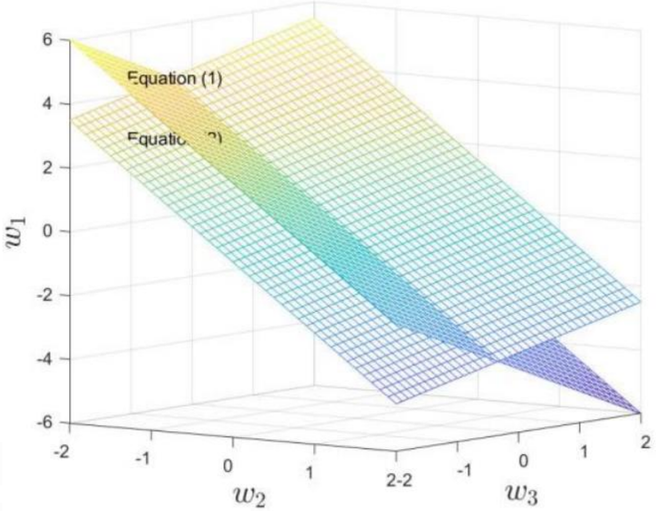
Given $\mathbf{w}^T\mathbf{X} = \mathbf{y}^T$ where

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

(a) What kind of system is this? (even-, over- or under-determined?)

(b) Is $\mathbf{X}$ invertible? Why?

(c) Solve for $\mathbf{w}$ if it is solvable.



Answer:

(a) This is an under-determined system (there are 3 unknowns with 2 equations).

(b) $\mathbf{X}$ is NOT invertible but $\mathbf{X}^T\mathbf{X}$ is. The determinant of $\mathbf{X}^T\mathbf{X} = 6\times21 - 9\times9 = 45$.

(c) A constrained solution (exact) is given by

$\hat{\mathbf{w}}^T = (\mathbf{X}\mathbf{a})^T$ (The 3-dimensional vector $\mathbf{w}$ can be constrained by projecting $\mathbf{X}$ onto a 2-dimensional vector $\mathbf{a}$)

$$= \mathbf{a}^T\mathbf{X}^T$$

$$= \mathbf{y}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \quad = [0 \quad 1]\begin{bmatrix} 0.4667 & -0.2 \\ -0.2 & 0.1333 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & -1 \end{bmatrix}$$

$$= [0.0667 \quad 0.1333 \quad -0.3333].$$

Q6c) Extended Working...

- From Lecture Notes L4

Under-determined system: $m < d$ in $\mathbf{Xw} = \mathbf{y}$,

For under-det systems, there are many / ∞ solutions. **w** represents this set of ∞ solns.

- Assume that there is a solution for **w** such that

$$\mathbf{w} = \mathbf{X}^T \cdot \mathbf{a}$$

1) Right now X is not a square matrix (m<d). By incorporating X^T, we can form $X.X^T$ which will be a square matrix that can inverted!

2) **a** here is then just an unknown variable, which is a subset of **w**.

- From Lecture Notes L4

Q6c)

Unknown that we want to solve for

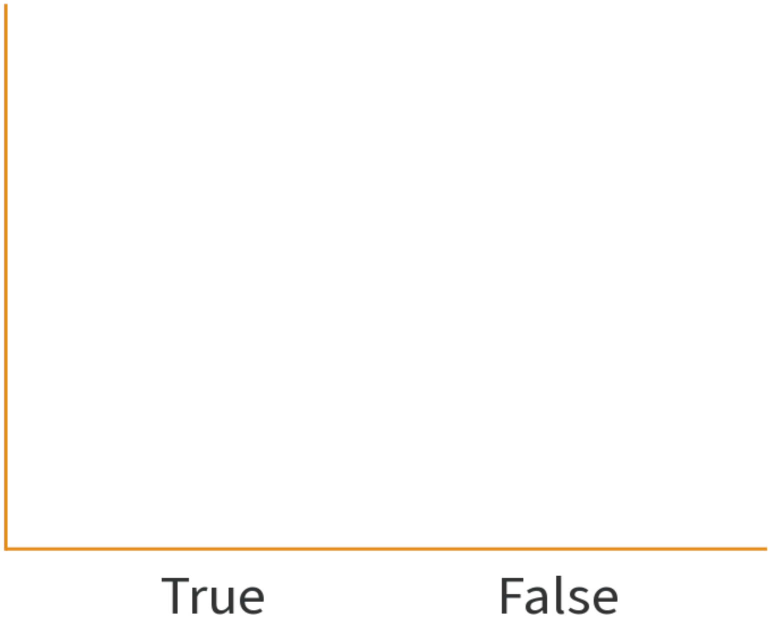
Some other unknown subset solution of **w.** Let's call it b.

Note that X^T is now on the right because we want to substitute this back into the original equation to form $X^T.X$.

NUS National University of Singapore

The following matrix has a left inverse.

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Q8 Which of the following is/are true about matrix A below? There could be more than one answer.