Solving Linear Equations

Case 2: Example - I

$$\begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.5 \\ 5 \end{bmatrix}$$

- $m = 3, n = 2$
- Using the optimization concept,

$$\boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -0.5 \\ 5 \end{bmatrix}$$

Case 2: Example continued

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.2 & -0.6 \\ -0.6 & 2.8 \end{bmatrix} \begin{bmatrix} 15 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$

- Thus, the solution for the given example is $(x_1, x_2) = (0,5)$
- Substituting in the equation shows

$$\begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} \neq \begin{bmatrix} 1 \\ -0.5 \\ 5 \end{bmatrix}$$

Case 2: Example

$$\begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

- $m = 3, n = 2$
- Using the optimization concept,

$$\boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

Case 2: Example continued

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.2 & -0.6 \\ -0.6 & 2.8 \end{bmatrix} \begin{bmatrix} 20 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

- Thus, the solution for the given example is $(x_1, x_2) = (1,2)$
- Substituting in the equation shows

$$\begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

R Code

```
A=matrix(c(1,2,3,0,0,1),ncol=2, byrow=F)
b=matrix(c(1,2,5),ncol=2, byrow=F)
x=inv(t(A)%*%A)%*%t(A)%*%b
x
```

Console output

```
> x=inv(t(A)%*%A)%*%t(A)%*%b
> x
     [,1]
[1,]    1
[2,]    2
>
```

Data science for Engineers

Case 3: $m < n$

- This case addresses the problem of more attributes or variables than equations
- Since the number of attributes is greater than the number of equations, one can obtain multiple solutions for the attributes
- This is termed as an infinite-solution case
- How does one choose a single solution from the set of infinite possible solutions?

Linear Algebra

Case 3: An optimization perspective

- Pose the following optimization problem

$$\min_x \left(\frac{1}{2}x^T x\right) \text{ s.t. } Ax = b$$

- Define a Lagrangian function $f(x, \lambda)$

$$\min \left[f(x, \lambda) = \frac{1}{2}x^T x + \lambda^T (Ax - b) \right]$$

- Differentiating the Lagrangian with respect to x, and setting to zero

$$x + A^T \lambda = 0$$

Data science for Engineers

Case 3: An optimization perspective

$$x = -A^T \lambda$$

Pre-multiplying by A

$$Ax = b = -AA^T \lambda$$

Thus we obtain $\lambda = -(AA^T)^{-1} b$ assuming that all the rows are linearly independent

$$\boldsymbol{x = -A^T \lambda = A^T \left(AA^T\right)^{-1} b}$$

Linear Algebra

Case 3: Example

$$\begin{bmatrix}1 & 2 & 3\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}$$

- $m = 2, n = 3$
- Using the optimization concept,

$$\boldsymbol{x} = \boldsymbol{A}^{\boldsymbol{T}}\left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}\right)^{-1}\boldsymbol{b}$$

$$x = \begin{bmatrix}1 & 0\\2 & 0\\3 & 1\end{bmatrix}\left(\begin{bmatrix}1 & 2 & 3\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\2 & 0\\3 & 1\end{bmatrix}\right)^{-1}\begin{bmatrix}2\\1\end{bmatrix}$$

Case 3: Example

$$x = \begin{bmatrix}1 & 0\\2 & 0\\3 & 1\end{bmatrix}\begin{bmatrix}14 & 3\\3 & 1\end{bmatrix}^{-1}\begin{bmatrix}2\\1\end{bmatrix}$$

$$x = \begin{bmatrix}1 & 0\\2 & 0\\3 & 1\end{bmatrix}\begin{bmatrix}-0.2\\1.6\end{bmatrix}$$

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-0.2\\-0.4\\1\end{bmatrix}$$

- The solution for the given example is $(x_1, x_2, x_3) = (-0.2, -0.4, 1)$

R Code

```
A=matrix(c(1,0,2,0,3,1),ncol=3)
b=c(2,1)
library(MASS)
x = t(A)%*%inv(A%*%t(A)) %*%b
x
```

Console output

```
A=matrix(c(1,0,2,0,3,1),ncol=3, byrow=F)
b=c(2,1)
x = t(A)%*%inv(A%*%t(A)) %*%b
x
```

Case 3: Example

$$\begin{bmatrix}1 & 2 & 3\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}2\\ 1\end{bmatrix}$$

- The solution for the given example is $(x_1, x_2, x_3) = (-0.2, -0.4, 1)$
- Verify this is a solution that satisfies the original equation
- This also turns out to be minimum norm solution

Data science for Engineers

Generalization

- The described cases cover all the scenarios one might encounter while solving linear equations
- Is there any form in which the results obtained for cases 1, 2 and 3 can be generalized ?
- The concept we used to generalize the solutions is called as Moore-Penrose pseudo-inverse of a matrix
- The pseudo inverse is used as follows

$$Ax = b$$

The solution becomes

$$x = A^{+}b$$

- Singular Value Decomposition can be used to calculate the pseudo inverse or the generalized inverse (A^{+})

Linear Algebra

Two examples revisited

Example 2

R Code
```
A=matrix(c(1,2,3,0,0,1),ncol=2, byrow=F)
b=matrix(c(1,2,5),ncol=1, byrow=F)
library(MASS)
x= ginv(A)%*%b
```

Solution

```
> x [,1]
[1,] 1
[2,] 2
```

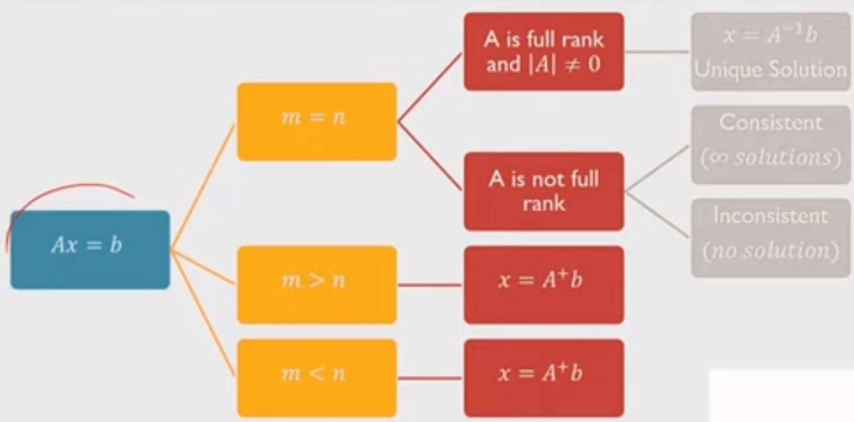
Example 3

R Code
```
A=matrix(c(1,0,2,0,3,1),ncol=3, byrow=F)
b=c(2,1)
library(MASS)
x = ginv(A)%*%b
```

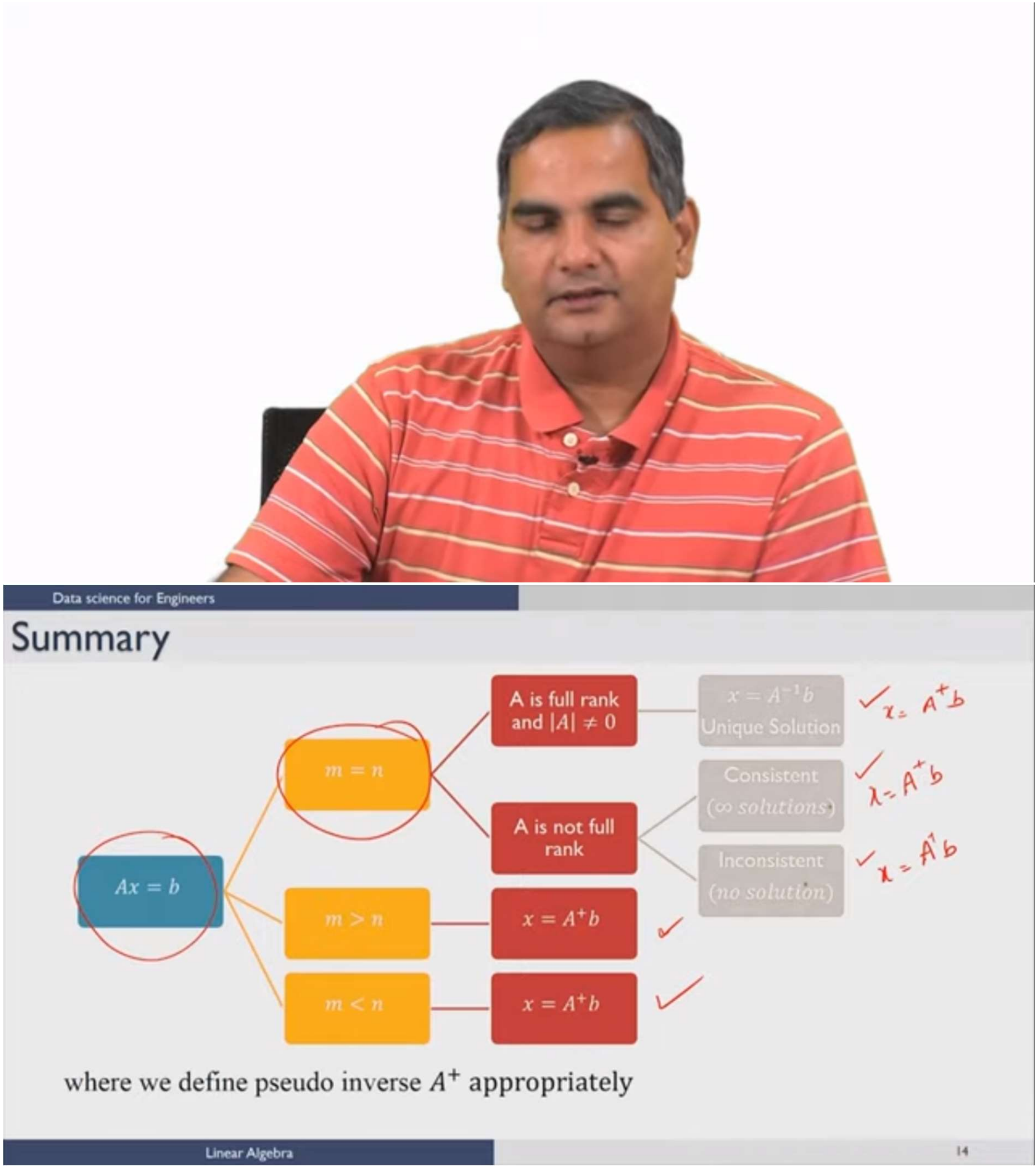
Solution

```
> x
     [,1]
[1,] -0.2
[2,] -0.4
[3,]  1.0
```

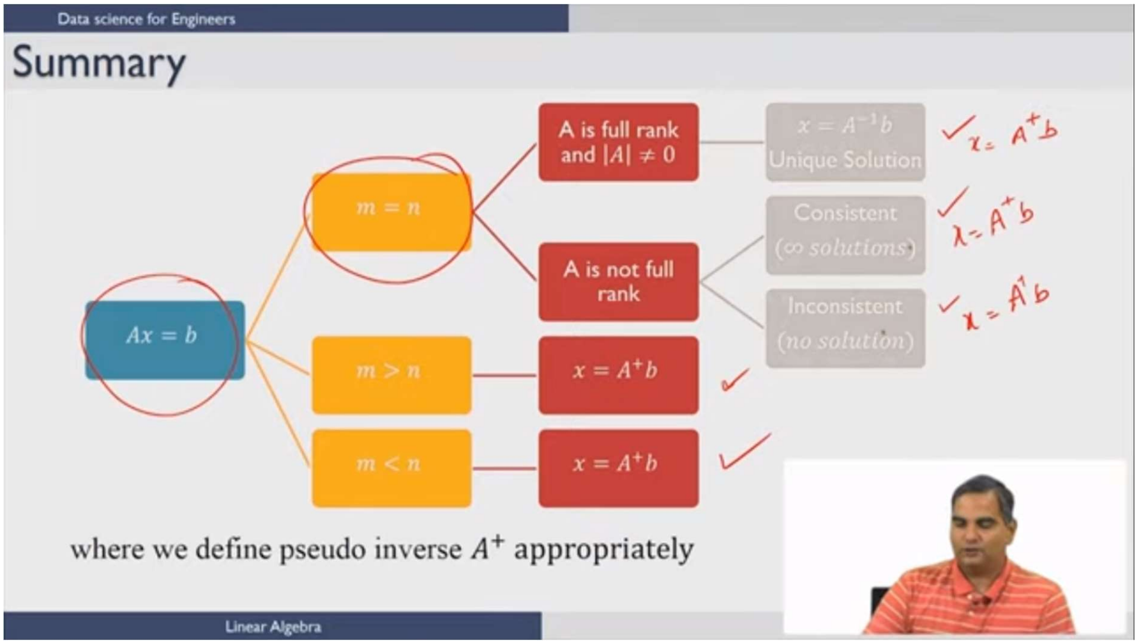

Summary



where we define pseudo inverse A^+ appropriately



Data science for Engineers
Summary
$Ax = b$
$m = n$
$m > n$
$m < n$
A is full rank and $|A| \neq 0$
A is not full rank
$x = A^{-1}b$ Unique Solution
Consistent (∞ solutions)
Inconsistent (no solution)
$x = A^{+}b$
$x = A^{+}b$
$x = A^{+}b$
$x = A^{+}b$
$x = A^{+}b$
where we define pseudo inverse A^{+} appropriately
Linear Algebra
14



Data science for Engineers
Summary
$Ax = b$
$m = n$
$m > n$
$m < n$
A is full rank and $|A| \neq 0$
A is not full rank
$x = A^{+}b$
$x = A^{+}b$
$x = A^{-1}b$ Unique Solution
Consistent (∞ solutions)
Inconsistent (no solution)
$x = A^{+}b$
$x = A^{+}b$
$x = A^{+}b$
where we define pseudo inverse A^{+} appropriately
Linear Algebra