EE2211 Lecture 4 (Vincent Tan, vtan@nus.edu.sg).

Reading: Lecture Slides Lec_4.pdf + Chapter 4 of EE2211 book.

About Me

- Associate Professor of Mathematics and ECE
- Joined NUS in 2014
- Undergraduate in Information Engineering from Cambridge University 2005
- Ph.D. in Electrical Engineering and Computer Science (EECS) from MIT in 2011
- Research interests in Information Theory, Signal Processing, and Machine Learning
- I teach probability, stochastic processes, information theory, machine learning and mathematical analysis at NUS.

Office hours will be announced soon.

Today's Agenda: i) Review of Linear Algebra concepts.

ii) Systems of Linear Equations.

iii) Python Demos.

Reminder about linear algebraic notions.

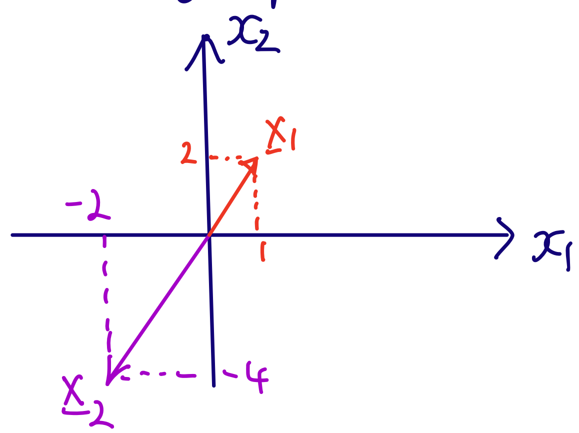
Def: A set of vectors $\{\underline{x}_1, \ldots, \underline{x}_k\}$, $\underline{x}_i \in \mathbb{R}^d$ (d-dimensional) is linearly independent if

$$\beta_1 \underline{x}_1 + \beta_2 \underline{x}_2 + \cdots + \beta_k \underline{x}_k = \underset{\in \mathbb{R}^d}{\underline{0}} \Rightarrow \beta_1 = \beta_2 = \cdots = \beta_k = 0.$$

Eg: $\underline{x}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ & $\underline{x}_2 = \begin{pmatrix} -2 \\ -4 \end{pmatrix}$. Are these vectors LI?

Consider $\quad \beta_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \beta_2 \begin{pmatrix} -2 \\ -4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

$\beta_1 = 2,\ \beta_2 = +1 \Rightarrow$ The set of vectors $\{x_1, x_2\}$ is not LI $\Rightarrow$ They are linearly dependent (LD)



Eg: $\underline{x}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ & $\underline{x}_2 = \begin{pmatrix} -2 \\ -3 \end{pmatrix}$. Are these vectors LI?

Consider $\quad \beta_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \beta_2 \begin{pmatrix} -2 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. — (*)

1st comp: $\quad \beta_1 - 2\beta_2 = 0 \Rightarrow \beta_1 = 2\beta_2$

2nd comp: $\quad 2\beta_1 - 3\beta_2 = 0 \quad 2(2\beta_2) - 3\beta_2 = 0 \Rightarrow \beta_2 = 0$

$\Rightarrow \beta_1 = 0$

Hence the only sol^{n} to (*) is $\beta_1 = \beta_2 = 0 \Rightarrow x_1, x_2$ are LI.

Def: Rank of a matrix A = # of pivots in its RREF.

Eg: $\text{rank}(A) = 3$

$$A = \begin{bmatrix} 2 & -2 & 4 & -2 \\ 2 & 1 & 10 & 7 \\ -4 & 4 & -8 & 4 \\ 4 & -1 & 14 & 6 \end{bmatrix} \overset{\text{ERO}}{\sim} \begin{bmatrix} 2 & -2 & 4 & -2 \\ 0 & 3 & 6 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 6 & 10 \end{bmatrix} \begin{array}{l} \\ R_2 - R_1 \\ R_3 + 2R_1 \\ R_4 - 2R_1 \end{array}$$

$$\sim A_{REF} = \begin{bmatrix} 2 & -2 & 4 & -2 \\ 0 & 3 & 6 & 9 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} R_4 - R_2 \Big\} \text{ swap } R_3 \,\&\, R_4$$

$$\sim A_{RREF} = \begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\text{rank}(A) = \text{rank}(A_{RREF}) = 3.$

Rank of A = # of linearly indep columns/rows of A.

$$A = \begin{bmatrix} 2 & -2 & 4 & -2 \\ 2 & 1 & 10 & 7 \\ -4 & 4 & -8 & 4 \\ 4 & -1 & 14 & 6 \end{bmatrix} \begin{array}{l} \to R_1 \\ \to R_2 \\ \to R_3 \\ \to R_4 \end{array} \qquad R_3 = -2R_1$$

Look at R_1, R_2, R_4. Check that these 3 rows are LI then $\text{rank}(A) = 3$.

Nature of Solutions to Linear Systems.

X : design matrix $\mathbb{R}^{m \times d}$ ($m \times d$ matrix)

y : target vector $\mathbb{R}^m$ (m-dim vector)

$$\begin{array}{cc} & \longleftarrow d \text{ columns} \longrightarrow \\ m \text{ rows} & \underbrace{\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & & \vdots \\ x_{m,1} & x_{m,2} & & x_{m,d} \end{bmatrix}}_{X \,:\, \text{given}} \end{array} \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_d \end{bmatrix}}_{\substack{\underline{w} \in \mathbb{R}^d \\ \text{unknown.}}} = \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{\underline{y} \,:\, \text{given}}$$

E.g. m students, d: characteristics / attributes / features.

$x_{i,j} = j^{th}$ characteristic of student $1 \le i \le m$.

e.g.: $x_{i,1}$ = # of hours student i studies for EE2211

$x_{i,2}$ = # of math classes student i has taken.

$$\begin{bmatrix} x_{11} & x_{12} \\ \vdots & \\ & \\ x_{m1} & x_{m,2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \longleftarrow \text{Exam mark of student } i.$$

For a linear system $X\underline{w} = \underline{y}$, we have

① unique solution $\underline{w}$

② no solution

③ infinitely many solutions. $\underline{w}$

$X \in \mathbb{R}^{m \times d}$, $\underline{w} \in \mathbb{R}^d$

$y \in \mathbb{R}^m$.

m rows ↓ ↙ d+1 columns.

Def: Augmented matrix $\tilde{X} = [X \;\; y] \in \mathbb{R}^{m \times (d+1)}$

$$\tilde{X} = \underbrace{\left[\begin{array}{cccc|c} \multicolumn{4}{c}{\overbrace{\hphantom{x_{1,1} \quad x_{1,2} \quad \cdots \quad x_{1,d}}}^{d}} & \\ x_{1,1} & x_{1,2} & \cdots & x_{1,d} & y_1 \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} & y_2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{m,1} & x_{m,2} & & x_{m,d} & y_m \end{array}\right]}_{d+1 \text{ cols}} \; \updownarrow m \text{ rows}$$

Fact: i) $X\underline{w} = y$ has a unique solution if

$$\operatorname{rank}(X) = \operatorname{rank}(\tilde{X}) = d.$$

ii) $X\underline{w} = y$ has no solution if

$$\operatorname{rank}(X) < \operatorname{rank}(\tilde{X})$$

iii) $X\underline{w} = y$ has infinitely-many solutions if

$$\operatorname{rank}(X) = \operatorname{rank}(\tilde{X}) < d.$$

$\operatorname{rank}(A) \leq \min\{m, n\}$

$A \in \mathbb{R}^{m \times n}$

Example of case (i):

$$X = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad y = \begin{bmatrix} 4 \\ 10 \end{bmatrix} \qquad d = 2$$

$\operatorname{rank}(X) = 2 \quad \because \begin{pmatrix} 2 \\ 4 \end{pmatrix} \,\&\, \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ are LI

$$\operatorname{rank}(\tilde{X}) = \operatorname{rank}\left(\begin{bmatrix} 2 & 1 & 4 \\ 4 & 3 & 10 \end{bmatrix}\right) = 2$$

$\operatorname{rank}(\tilde{X}) \geq \operatorname{rank}(X)$

We are in case (i) $\Rightarrow$ There is a unique sol.

$\text{rank}(X) = \text{rank}(\tilde{X}) = d = 2.$

Example of Case (ii): Usual situation.

$$X = \begin{bmatrix} 2 & 1 \\ 4 & 3 \\ 5 & 6 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$d = 2$
$m = 3$
$m > d$: over-determined system

$\text{rank}(X) = 2$

$$\text{rank}(\tilde{X}) = \text{rank}\left(\begin{bmatrix} 2 & 1 & 1 \\ 4 & 3 & 2 \\ 5 & 6 & 3 \end{bmatrix}\right) = 3. \quad \text{(You check)}$$

$\text{rank}(X) < \text{rank}(\tilde{X}) \Rightarrow$ Case (ii) $\Rightarrow$ no solution.

Intuitively, we have more equations ($m = 3$) than unknowns ($d = 2$). Hence the system of eqns is overconstrained $\Rightarrow$ no solution.

Example of Case (i): Unusual situation.

$$X = \begin{bmatrix} 2 & 1 \\ 4 & 3 \\ 5 & 6 \end{bmatrix}, \quad y = \begin{bmatrix} 4 \\ 10 \\ 17 \end{bmatrix}$$

$d = 2$
$m = 3$
$m > d$: over-determined system

$\text{rank}(X) = 2$

$$\text{rank}(\tilde{X}) = \text{rank}\left(\begin{bmatrix} 2 & 1 & 4 \\ 4 & 3 & 10 \\ 5 & 6 & 17 \end{bmatrix}\right) = 2.$$

$$\begin{pmatrix} 4 \\ 10 \\ 17 \end{pmatrix} = 1 \cdot \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix}$$

The 3 columns of $\tilde{X}$ are Linearly dependent

$\text{rank}(X) = \text{rank}(\tilde{X}) = d = 2 \Rightarrow$ Case (i): unique sol^{n}.

Example of Case (iii): Usual Situation

$$X = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 5 \end{bmatrix}, \quad y = \begin{bmatrix} 10 \\ 7 \end{bmatrix}$$

$m = 2$
$d = 3$

$m < d$ underdetermined system.

$$\text{rank}(X) = \text{rank}(\tilde{X}) = \text{rank}\left(\begin{bmatrix} 2 & 1 & 3 & 10 \\ 4 & 2 & 5 & 7 \end{bmatrix}\right) = 2$$

So $\text{rank}(X) = \text{rank}(\tilde{X}) < d = 3$.

$\Rightarrow$ Case (iii). Infinitely-many sol^{n}.

Intuition: We have more unknowns ($d=3$) than equations ($m=2$) so not enough constraints for unique sol^{n}.

Can an under-det system have no solution?
$m=2 \quad d=3$

Unusual: $X = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 6 \end{bmatrix} \qquad y = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$

$\text{rank}(X) = 1 \qquad \underbrace{\text{rank}\left(\begin{bmatrix} 2 & 1 & 3 & 1 \\ 4 & 2 & 6 & 3 \end{bmatrix}\right)}_{\text{rank}(\tilde{X})\,=} = 2$

$\Rightarrow \text{rank}(X) < \text{rank}(\tilde{X})$

$\Rightarrow$ No solution! Even though the system is underdetermined.

$\Rightarrow$ Unusual scenario $\because$ # unknowns $(d=3)$ > # eqns $(m=2)$

So we expect that there are ∞-many sol$^{\underline{n}}$.

Solving a Linear System

Square, Even-determined System $m=d \qquad X\underline{w} = \underline{y}$

$$\underbrace{\begin{bmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & & \vdots \\ x_{m1} & & x_{md} \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_d \end{bmatrix}}_{\underline{w}} = \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{\underline{y}}$$

If X has full rank (i.e., $\underline{\text{rank}(X) = m = d}$)
then there is a unique solution.

Why? $\text{rank}(X) = \text{rank}(\tilde{X}) = d \qquad \tilde{X} \in \mathbb{R}^{d \times (1+d)}$

$\Rightarrow$ Case (i) $\Rightarrow$ $\exists$ unique sol$^{\underline{n}}$

$$\cancel{X^{-1}}\cancel{X}\,\underline{\hat{w}} = X^{-1}\underline{y} \Rightarrow \underline{\hat{w}} = X^{-1}\underline{y}.$$

The inverse of X, say X^{-1}, satisfies $XX^{-1} = X^{-1}X = I$
If X^{-1} exists, we say that X is invertible/non-singular/full rank/det $\neq 0$.

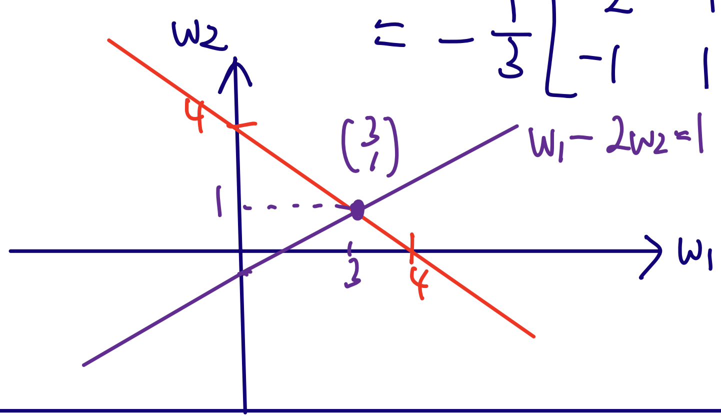
Ex: $\boxed{w_1 + w_2 = 4}$ $\boxed{w_1 - 2w_2 = 1}$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}}_{\underline{w}} = \underbrace{\begin{bmatrix} 4 \\ 1 \end{bmatrix}}_{\underline{y}}$$

Since X is non-singular $(\det(X) = 1(-2) - 1\cdot 1 \neq 0)$,

$$\hat{\underline{w}} = X^{-1}y = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

$$= -\frac{1}{3}\begin{bmatrix} -2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

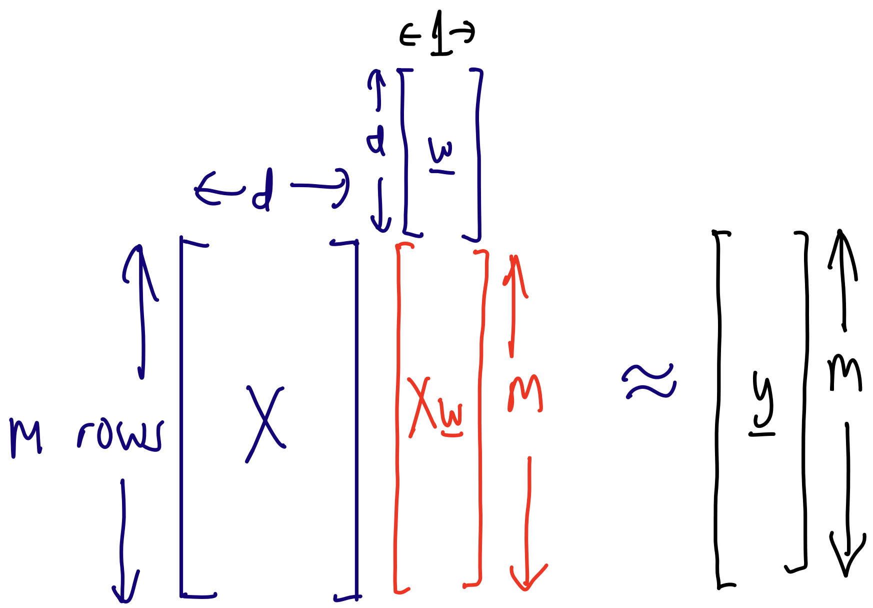


$w_1 - 2w_2 = 1$

Overdetermined

$m = \#$ equations $> d = \#$ unknowns.

- X tall matrix
- X: non-square $\Rightarrow$ not invertible
- Usual, general scenario $\Rightarrow$ there is <u>no</u> sol^{n}.
- However, we can find an approximate sol^{n}, called the least squares sol^{n}



Consider $X^{+} = (X^TX)^{-1}X^T$: left-inverse of X.

$$X^{+}X = (X^TX)^{-1}(X^TX) = I_d$$

$(X^TX)^{-1}$ exists if X has <u>full column rank</u>, i.e., $\#$ of LI columns $= d$.

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

$$\underbrace{(X^TX)^{-1}(X^TX)}_{I}\hat{w} = (X^TX)^{-1}X^Ty$$

$$\hat{w}_{LS} = (X^T X)^{-1} X^T y.$$

E.g. $X = \begin{bmatrix} 2 & 1 \\ 4 & 3 \\ 5 & 6 \end{bmatrix}$

Col-rank$(X) = 2 =$ # of col $\Rightarrow$ full col. rank.

$X = \begin{bmatrix} 2 & 4 \\ 4 & 8 \\ 5 & 10 \end{bmatrix}$

Col-rank$(X) = 1 <$ # of col $= 2 \Rightarrow$ not full col. rank.

Ex: $w_1 + w_2 = 1$, $w_1 - w_2 = 0$, $w_1 = 2$

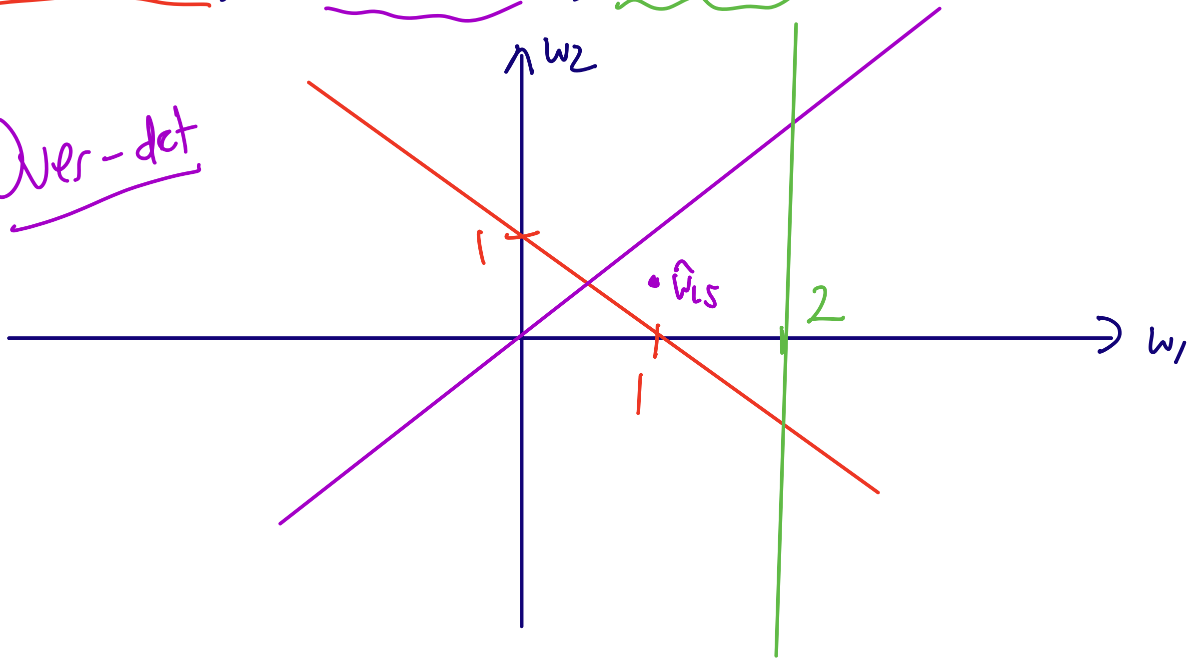
$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}}_{\underline{w}} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}}_{y}$$

$(X^T X)^{-1}$ exists $\because X$ has full col. rank.

$$\hat{w}_{LS} = (X^T X)^{-1} X^T y = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}.$$

$w_1 + w_2 = 1$, $w_1 - w_2 = 0$, $w_1 = 2$

Over-det



Underdetermined $m = \#\text{ eqns} < \#\text{ unknowns} = d$.

Usually we have infinitely many solutions.

X: fat

X: not-invertible.

- Among the multiple sol^{n}s, we want to find one special sol.
  $\Rightarrow$ least norm solution.

$$\overbrace{\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ \vdots & & & \\ x_{m1} & x_{m2} & \cdots & x_{md} \end{bmatrix}}^{d \text{ columns}}_{m \times d} \underbrace{\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix}}_{d} = \underbrace{\begin{bmatrix} Xw \end{bmatrix}}_{1} = \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{m}$$

Least norm solution to the linear system $X\underline{w}=y$ is

$$\hat{w}=X^{+}y$$

X^{+}: right-inverse $\qquad X^{+}= X^{T}(XX^{T})^{-1}$

$$XX^{+}=(XX^{T})(XX^{T})^{-1}=I$$

Least norm solution: $\quad \hat{w}_{LN}= X^{T}(XX^{T})^{-1}y.$

Right-inverse exists $\Leftrightarrow$ X has full row rank.

$\Leftrightarrow$ # of LI rows = # rows.

Consider: $\quad w_1+2w_2+3w_3=2, \quad w_1-2w_2+3w_3=1$

$$\begin{bmatrix}1 & 2 & 3\\ 1 & -2 & 3\end{bmatrix}\begin{bmatrix}w_1\\ w_2\\ w_3\end{bmatrix}=\begin{bmatrix}2\\ 1\end{bmatrix}$$

Least norm sol$^{\underline{n}}$: $\quad \hat{w}_{LN}= X^{T}(XX^{T})^{-1}y=\begin{bmatrix}0.15\\ 0.25\\ 0.45\end{bmatrix}$

$$\underline{\hat{w}}=\begin{bmatrix}0.15\\ 0.25\\ 0.45\end{bmatrix}+t\begin{bmatrix}-3\\ 0\\ 1\end{bmatrix}, \quad t\in\mathbb{R}$$

$\hat{w}_{LS}=(X^{T}X)^{-1}X^{T}y$ $\qquad$ $\hat{w}_{LN}=X^{T}(XX^{T})^{-1}y$

overdet $\qquad$ underdet