X — N rows, $K+1$ columns. Last column is all "1"

Y target vector

Find M $(K+1)\times 1$ vector so that

$$\|Y - XM\|^2 \text{ is minimized.}$$

This happens when

$$X^T(Y - XM) = 0$$

$$X^TY = X^TXM \qquad D \quad (K+1)\times(K+1)$$

$$D = X^TX$$

$$M = D^{-1}X^TY$$ provided D is an invertible matrix.

Remember:

D is invertible $\Leftrightarrow$ $\ker D = 0$ $\Leftrightarrow$ when $Dv = 0$ for some v, we have $v = 0$.

D is invertible $\Leftrightarrow$ when v is a nonzero vector, $Dv \neq 0$.

<u>Prop</u>: D is invertible if and only if the columns of X are linearly independent.

cols of $X \in \mathbb{R}^N$ — colspace of X has dim $K+1$

coldim $X = K+1$.

2-dim'l case.

$$X = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} \lambda \\ \vdots \\ \lambda \end{pmatrix}$$ columns are <u>dependent</u>

Higher dim'l case:

coldim$(X) = K+1$ means that the <u>features</u> are independent.

if this happens, features are redundant.

Proof: First suppose that cols of X are dependent.

$$X[:,1]m_1 + X[:,2]m_2 + \cdots + X[:,k+1]m_{k+1} = 0$$

but not all $m_i = 0$.

same as: $$X\begin{pmatrix} m_1 \\ \vdots \\ m_{k+1} \end{pmatrix} = 0.$$

$$\underline{X^T X}\begin{pmatrix} m_1 \\ \vdots \\ m_{k+1} \end{pmatrix} = 0$$

$$D\begin{pmatrix} m_1 \\ \vdots \\ m_{k+1} \end{pmatrix} = 0$$

so $\ker D$ contains a nonzero vector.

D is <u>NOT</u> invertible.

Suppose D NOT invertible. Then there is a vector $m = \begin{pmatrix} m_1 \\ \vdots \\ m_{k+1} \end{pmatrix} \neq 0$ so that $Dm = 0$.

$$X^T X m = 0.$$

$Xm \in$ colspace of X.

$\underset{k+1 \times N}{X^T} \cdot \underset{N \times 1 \text{ vector}}{Xm}$ is $k+1 \times 1$ vector.

entries of $X^T X m$ are the dot products of the rows of X^T with Xm.

Rows of X^T are the columns of X.

$X^T X m = 0$ means Xm is $\perp$ to all the columns of X.

$v \in$ colspace of X

$v = v_1 X[:,1] + v_2 X[:,2] + \dots + v_{k+1} X[:,k+1]$.

$XM \cdot v = \sum_{i=1}^{k+1} v_i \left(XM \cdot X[:,i] \right) = 0$

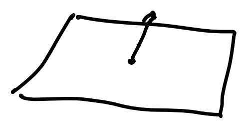
XM is $\perp$ to colspace of X

$XM \in \text{colspace}(X)$.

<u>Theorem</u>: if a vector $v \in W \subseteq V$ and

v is orthogonal to every elt of W

then $v = 0$



$X^T X m = 0$

$\Rightarrow XM = 0$

$\Rightarrow$ colspace(X) has dim $< k+1$

columns of X are dependent.