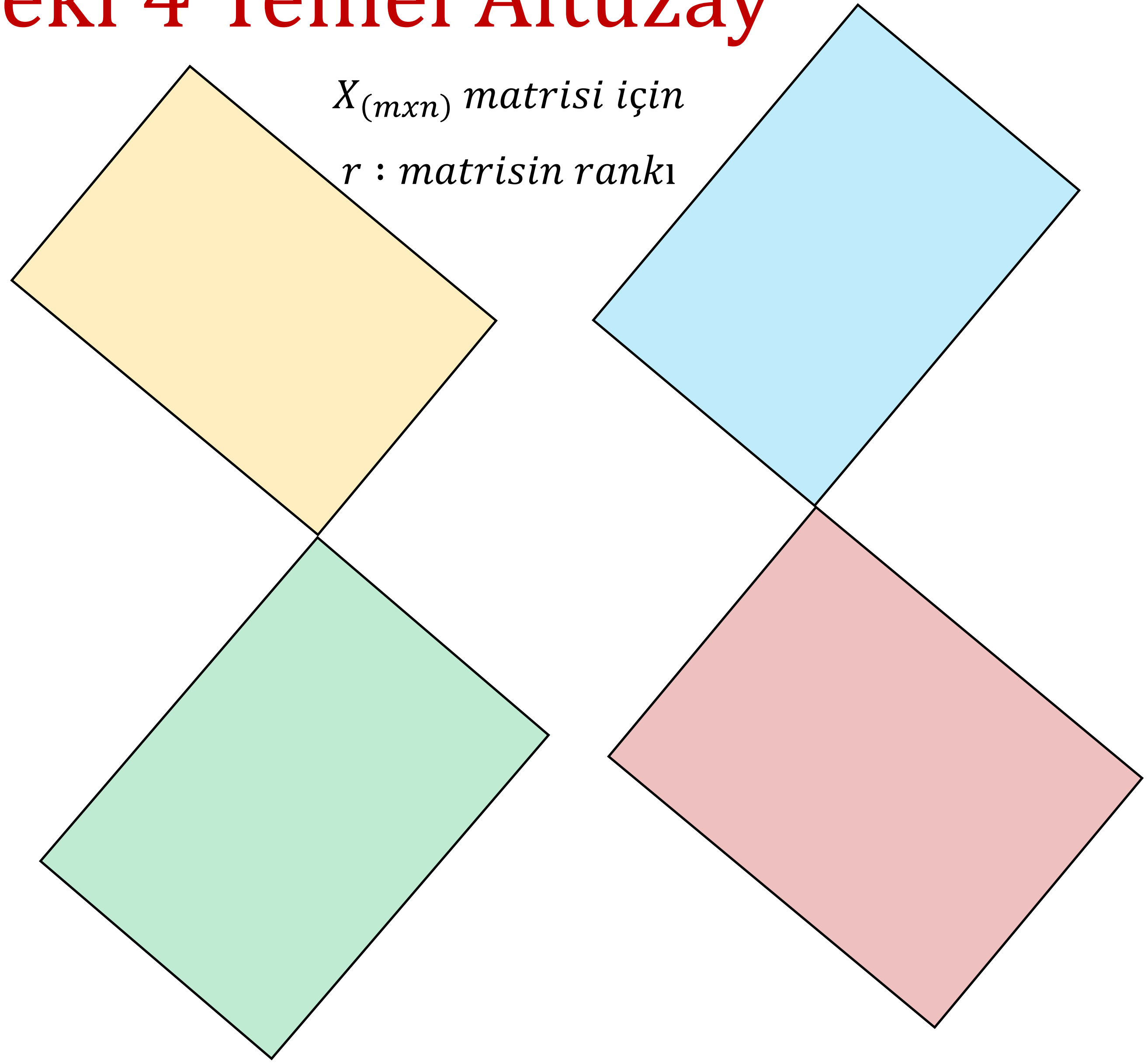
Lineer Regresyon : Normal Denklem

Lineer Cebirdeki 4 Temel Altuzay

$X_{(mxn)}$ *matrisi için*

r : *matrisin rankı*

$\mathcal{R}^n$ *boyutunda*

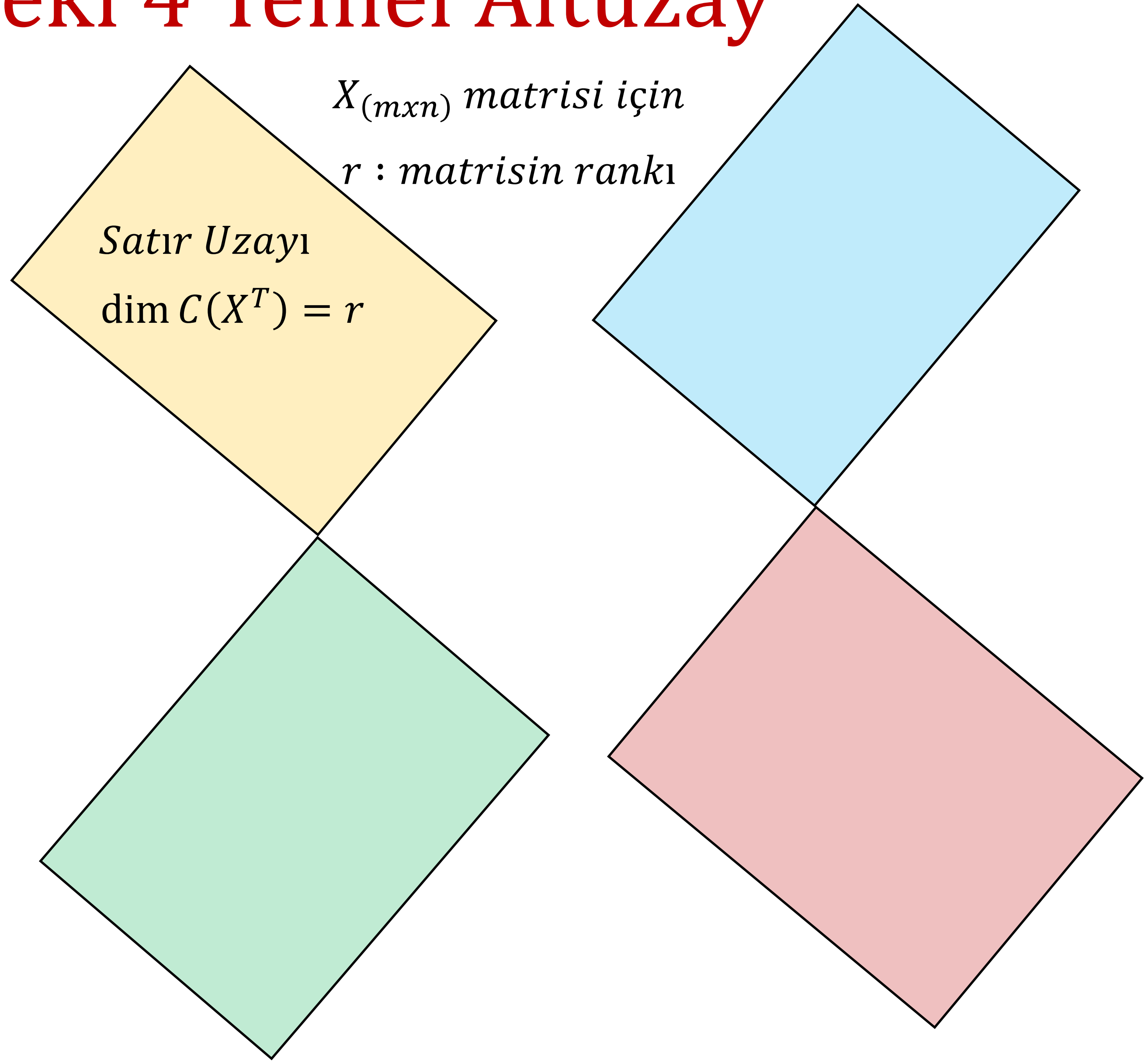


$\mathcal{R}^m$ *boyutunda*

Lineer Cebirdeki 4 Temel Altuzay

$X_{(mxn)}$ *matrisi için*

r : *matrisin rankı*



$\mathcal{R}^n$ *boyutunda*

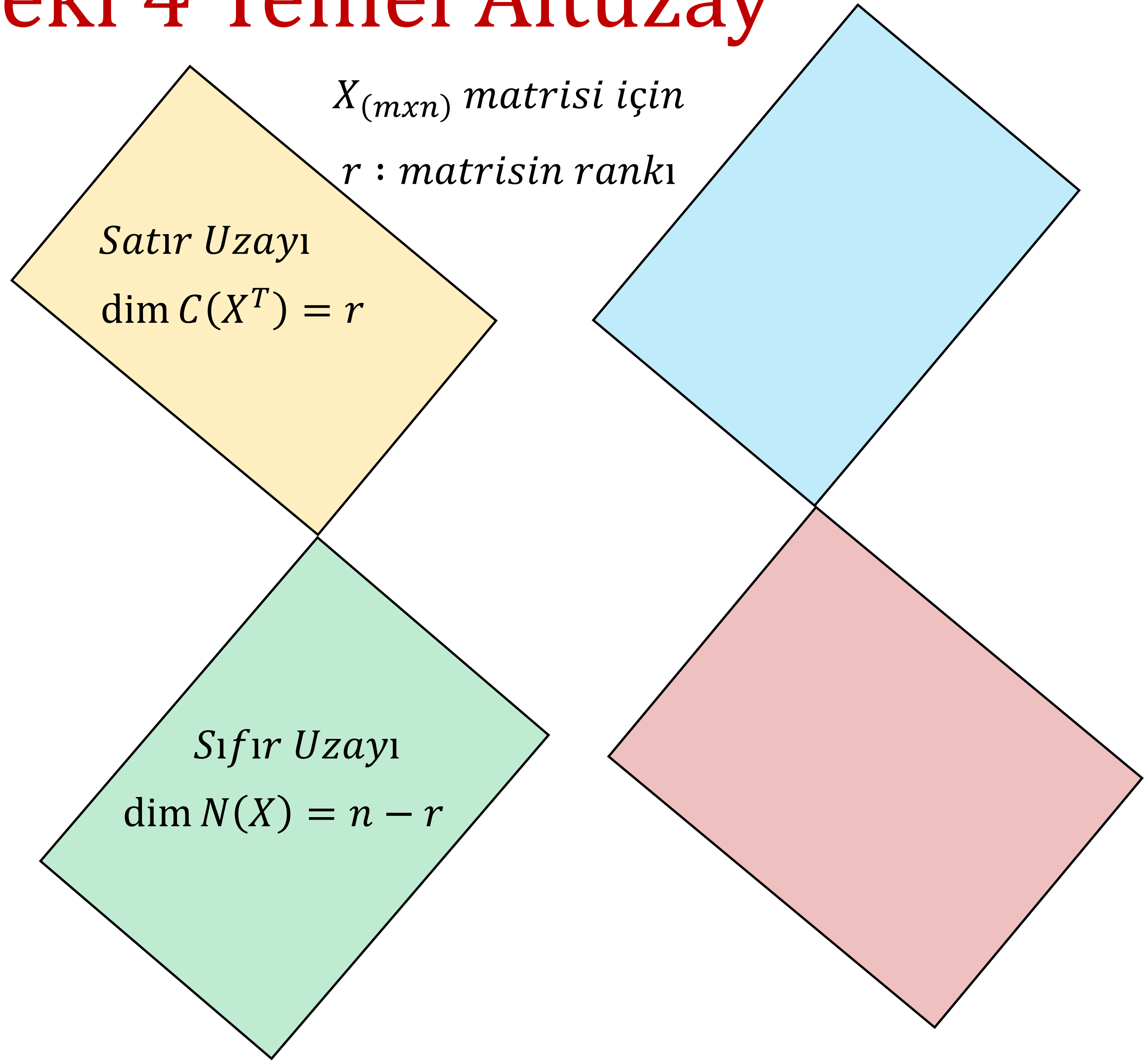
$\mathcal{R}^m$ *boyutunda*

Lineer Cebirdeki 4 Temel Altuzay

$X_{(mxn)}$ *matrisi için*

r : *matrisin rankı*

$\mathcal{R}^n$ *boyutunda*

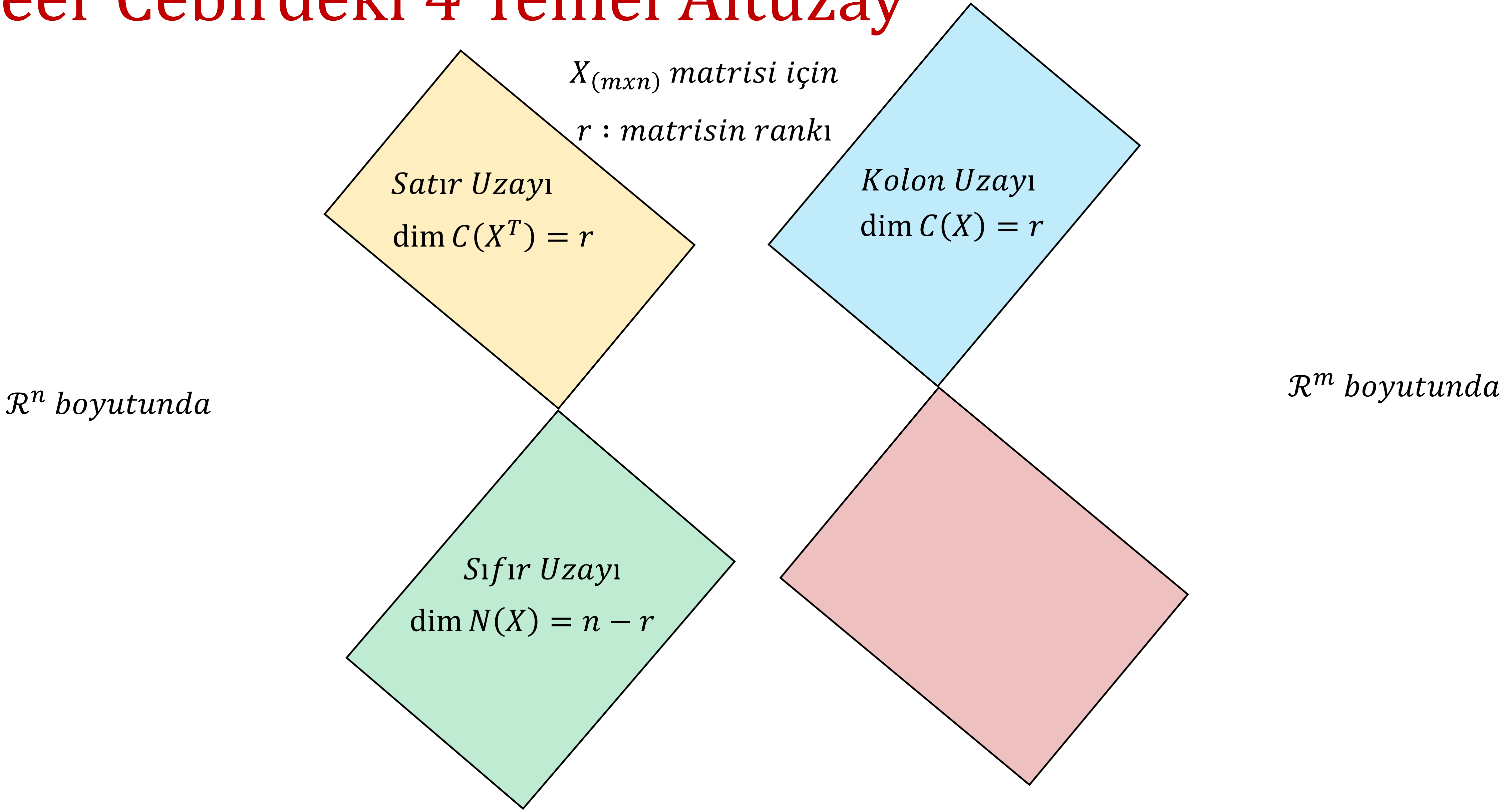


$\mathcal{R}^m$ *boyutunda*

Lineer Cebirdeki 4 Temel Altuzay

$X_{(mxn)}$ *matrisi için*

r : *matrisin rankı*



Lineer Cebirdeki 4 Temel Altuzay

$X_{(mxn)}$ *matrisi için*

r : *matrisin rankı*

Satır Uzayı

$\dim C(X^T) = r$

Kolon Uzayı

$\dim C(X) = r$

$\mathcal{R}^n$ *boyutunda*

$\mathcal{R}^m$ *boyutunda*

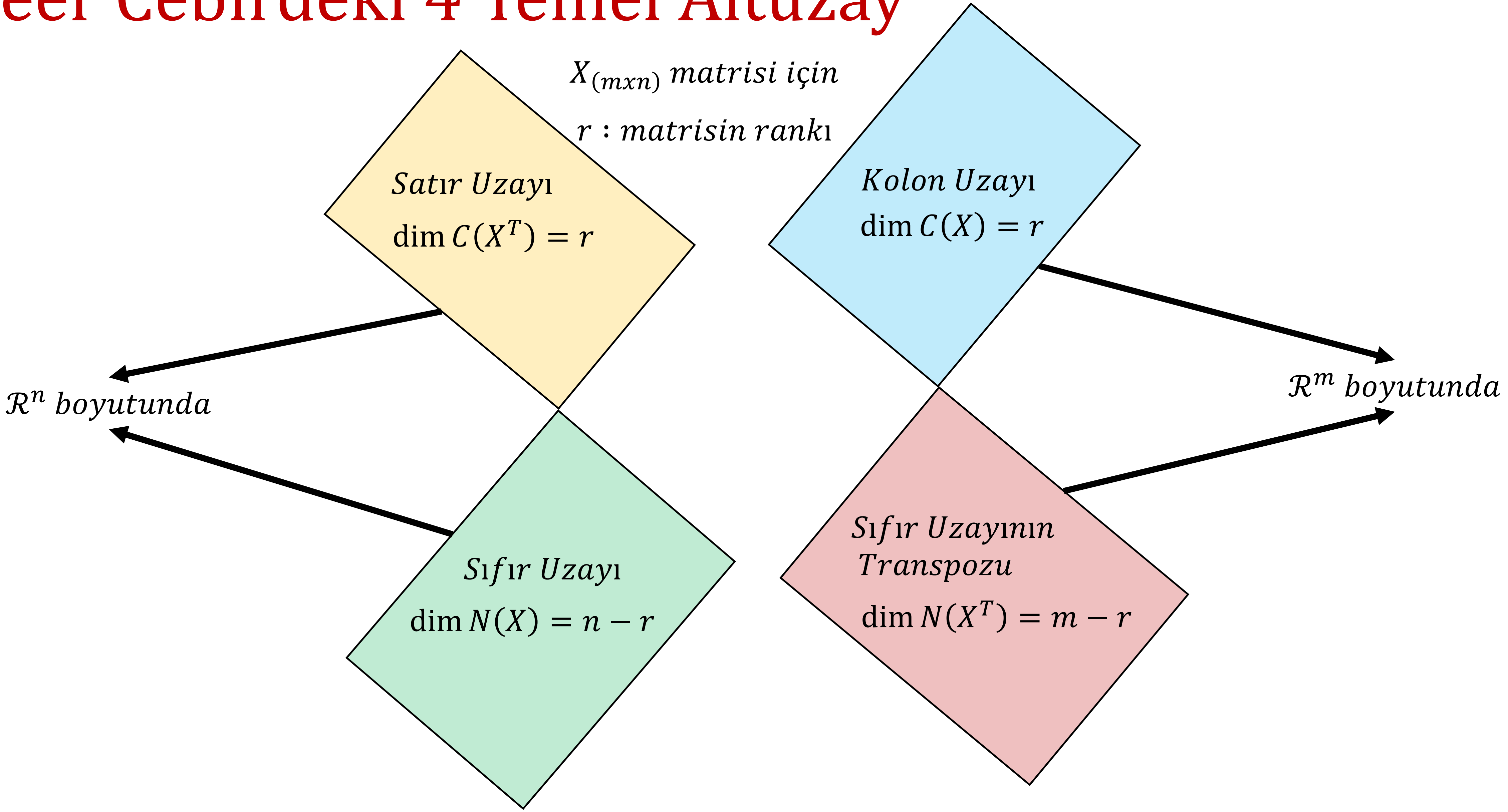
Sıfır Uzayı

$\dim N(X) = n - r$

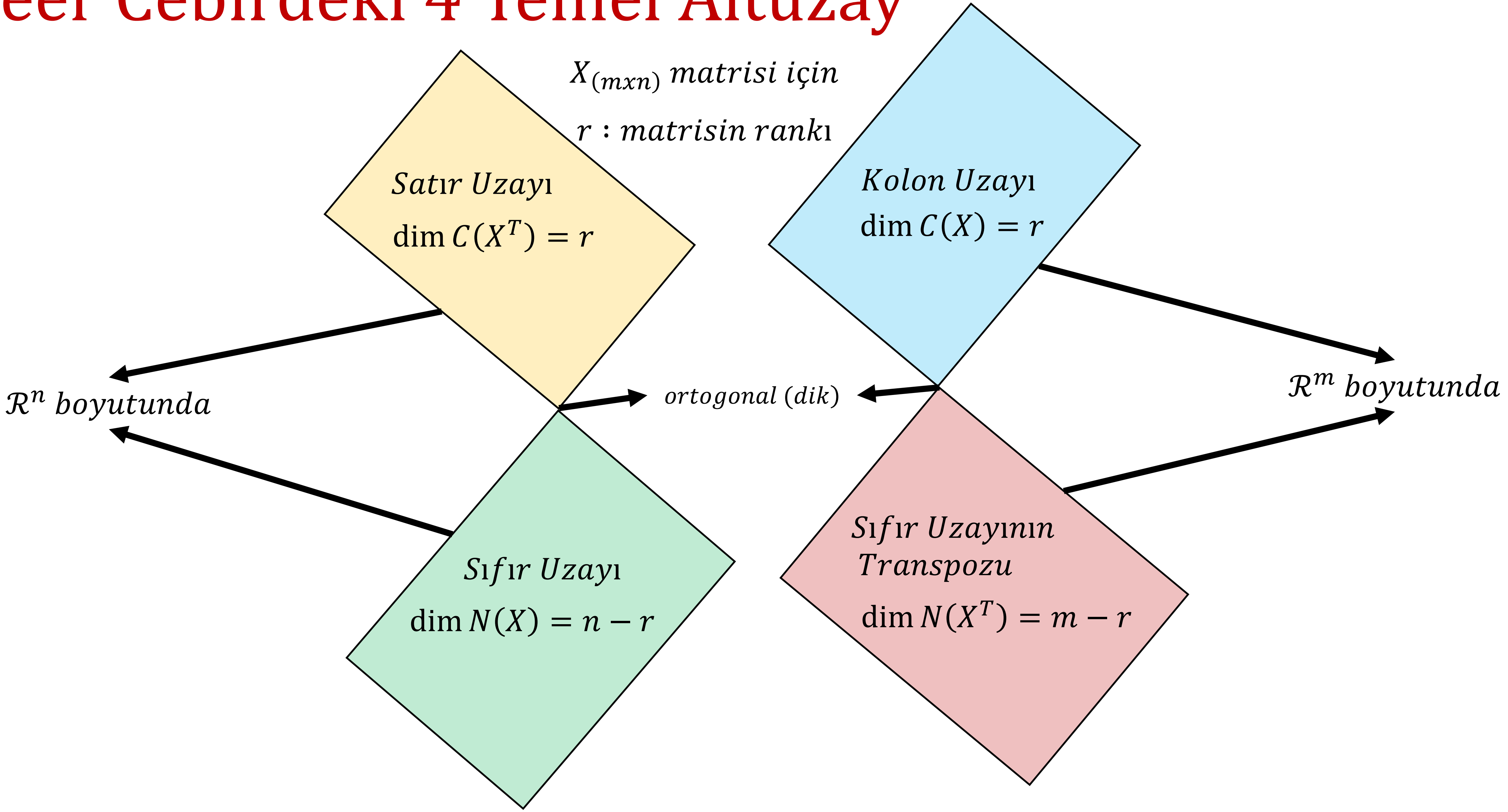
Sıfır Uzayının Transpozu

$\dim N(X^T) = m - r$

Lineer Cebirdeki 4 Temel Altuzay



Lineer Cebirdeki 4 Temel Altuzay



Kolon Uzayı (Column Space)

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & & \vdots & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & & \vdots & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = w_1 \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \end{bmatrix} + w_2 \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{m2} \end{bmatrix} + \cdots + w_n \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{mn} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Sıfır Uzayı (Null Space)

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & \vdots & & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & \vdots & & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = w_1 \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \end{bmatrix} + w_2 \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{m2} \end{bmatrix} + \cdots + w_n \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{mn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Satır Uzayı (Row Space)

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & & \vdots & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}_{nxm}^{T} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}_{mx1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{mx1}$$

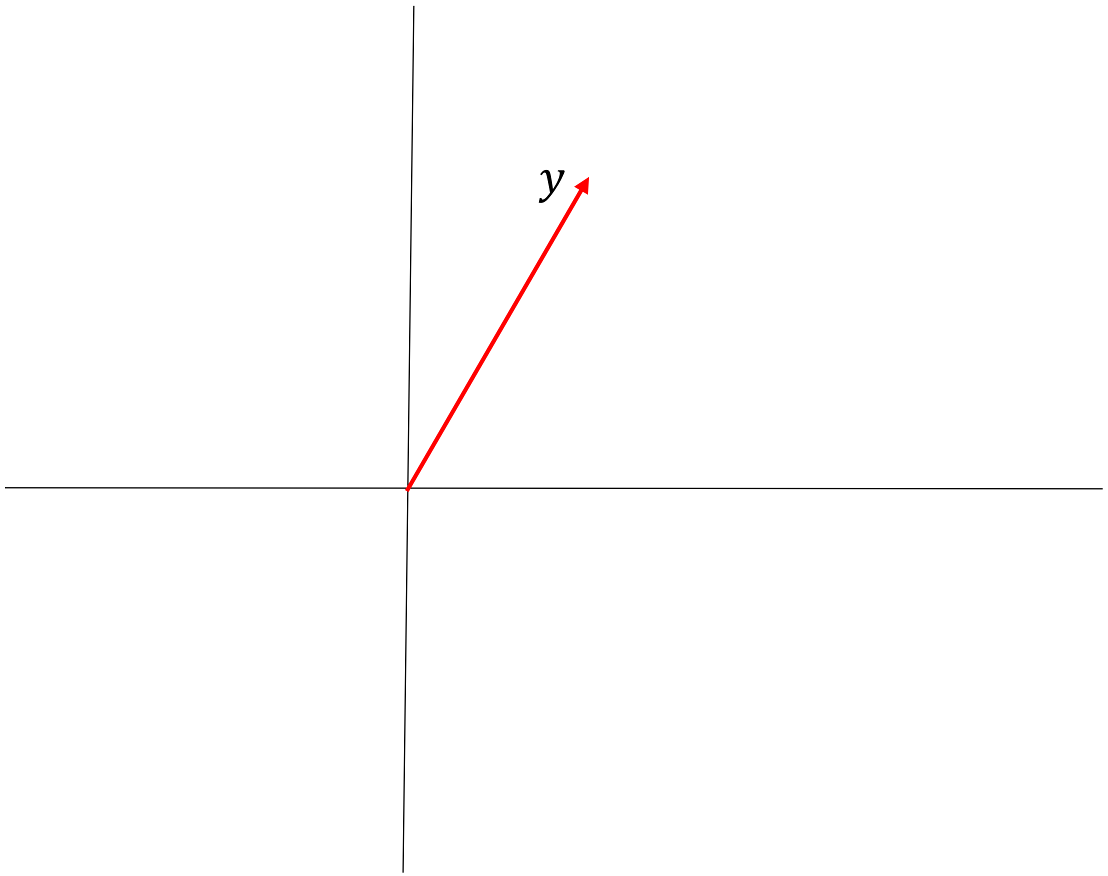
$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & & \vdots & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}_{nxm}^{T} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}_{mx1} = w_1 \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{bmatrix} + w_2 \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{bmatrix} + \cdots + w_m \begin{bmatrix} x_{m1} \\ x_{m2} \\ \vdots \\ x_{mn} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

X Transpozun Sıfır Uzayı

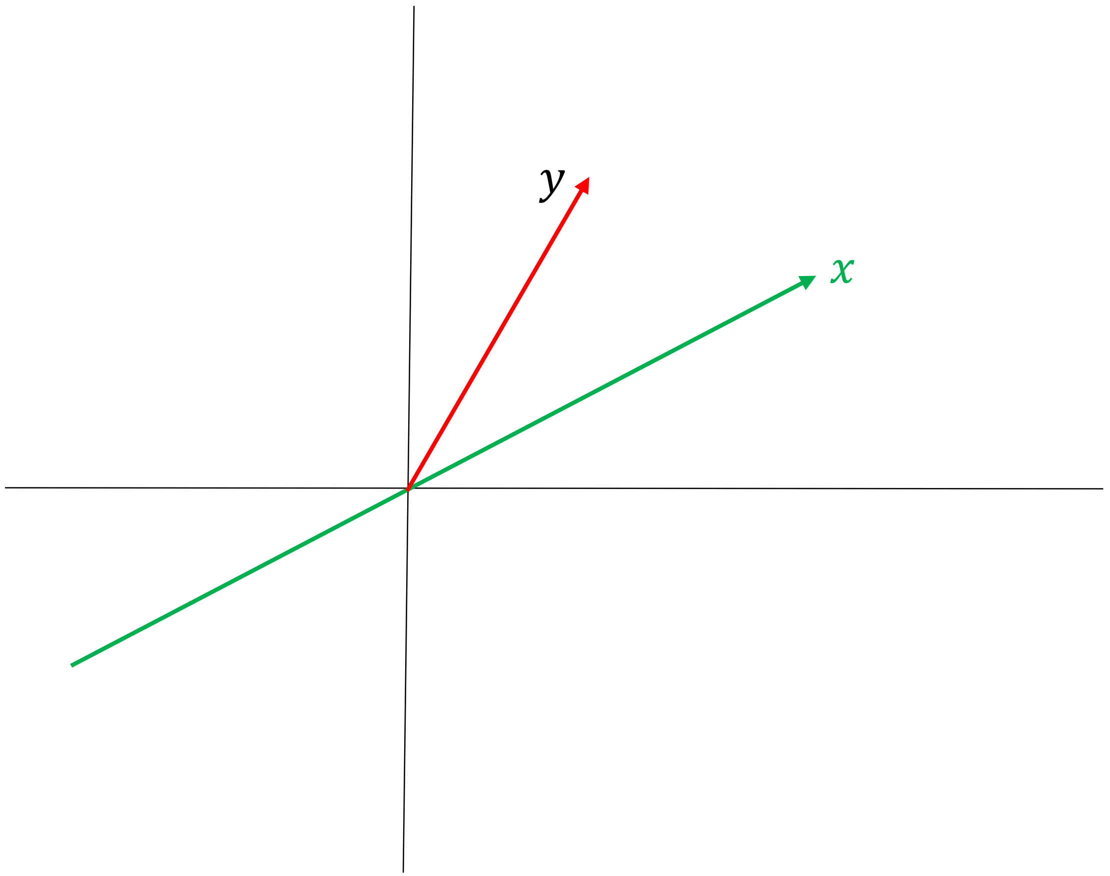
$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & \vdots & & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}_{nxm}^{T} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}_{mx1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{nx1}$$

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & \vdots & & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}_{nxm}^{T} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}_{mx1} = w_1 \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{bmatrix} + w_2 \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{bmatrix} + \cdots + w_m \begin{bmatrix} x_{m1} \\ x_{m2} \\ \vdots \\ x_{mn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

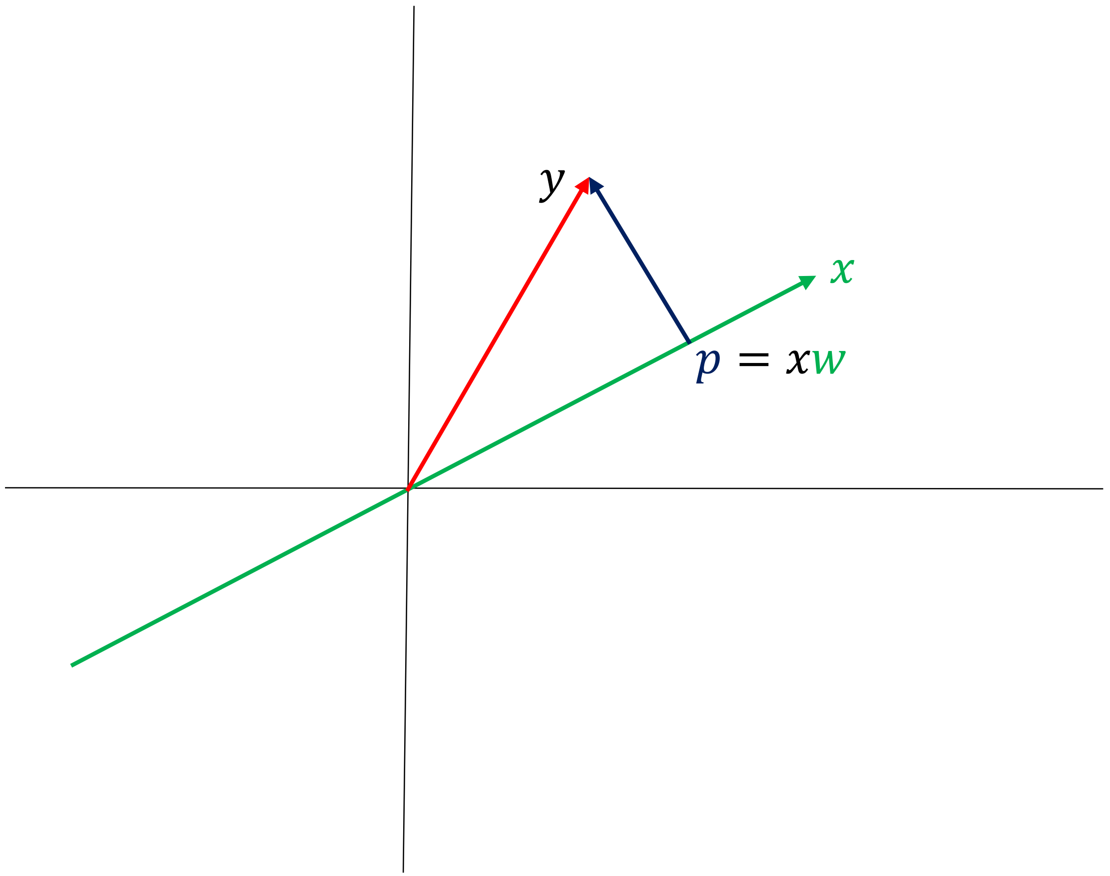
Projeksiyon ve Projeksiyon Matrisi



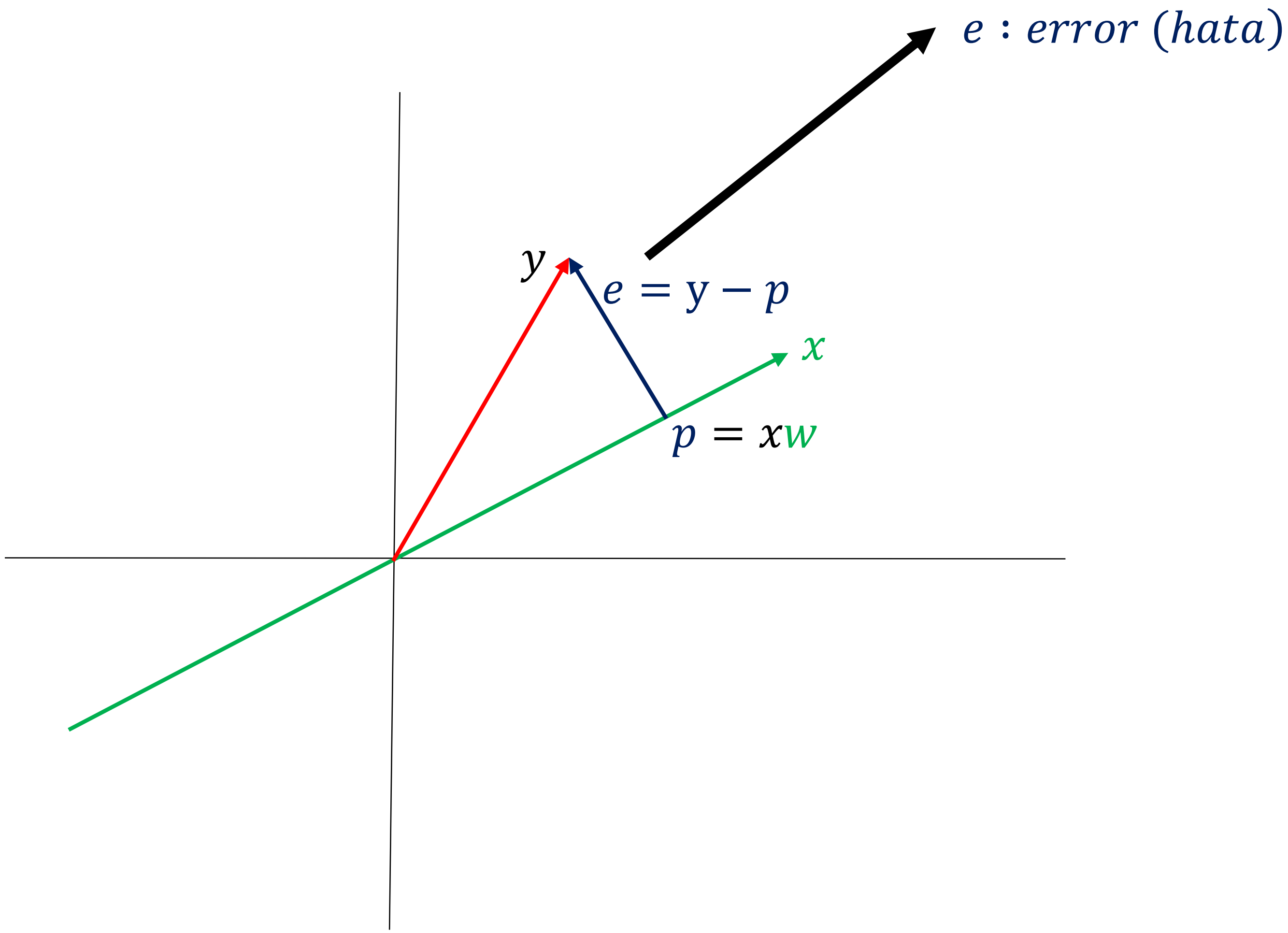
Projeksiyon ve Projeksiyon Matrisi



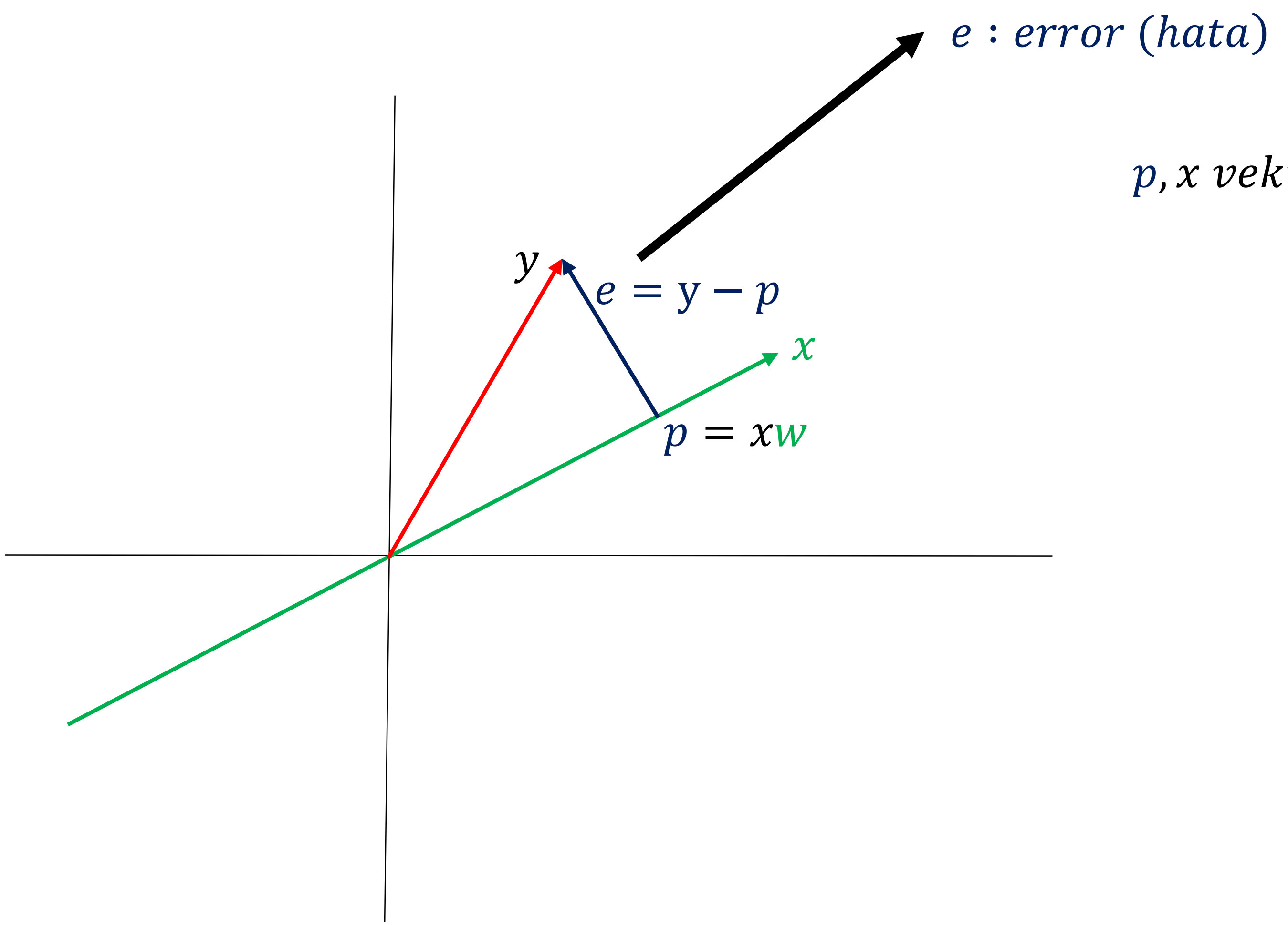
Projeksiyon ve Projeksiyon Matrisi



Projeksiyon ve Projeksiyon Matrisi

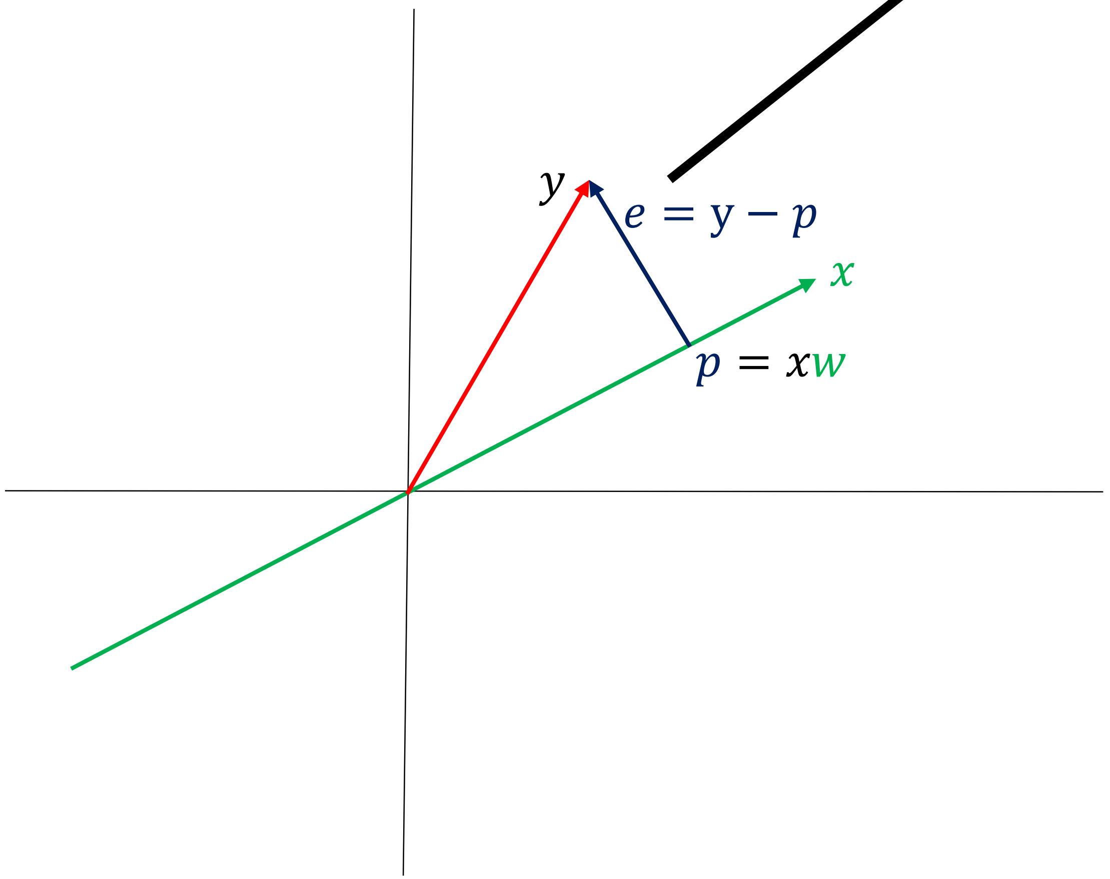


Projeksiyon ve Projeksiyon Matrisi



$p, x\ vektörünün\ herhangi\ bir\ katı : p = xw$

Projeksiyon ve Projeksiyon Matrisi

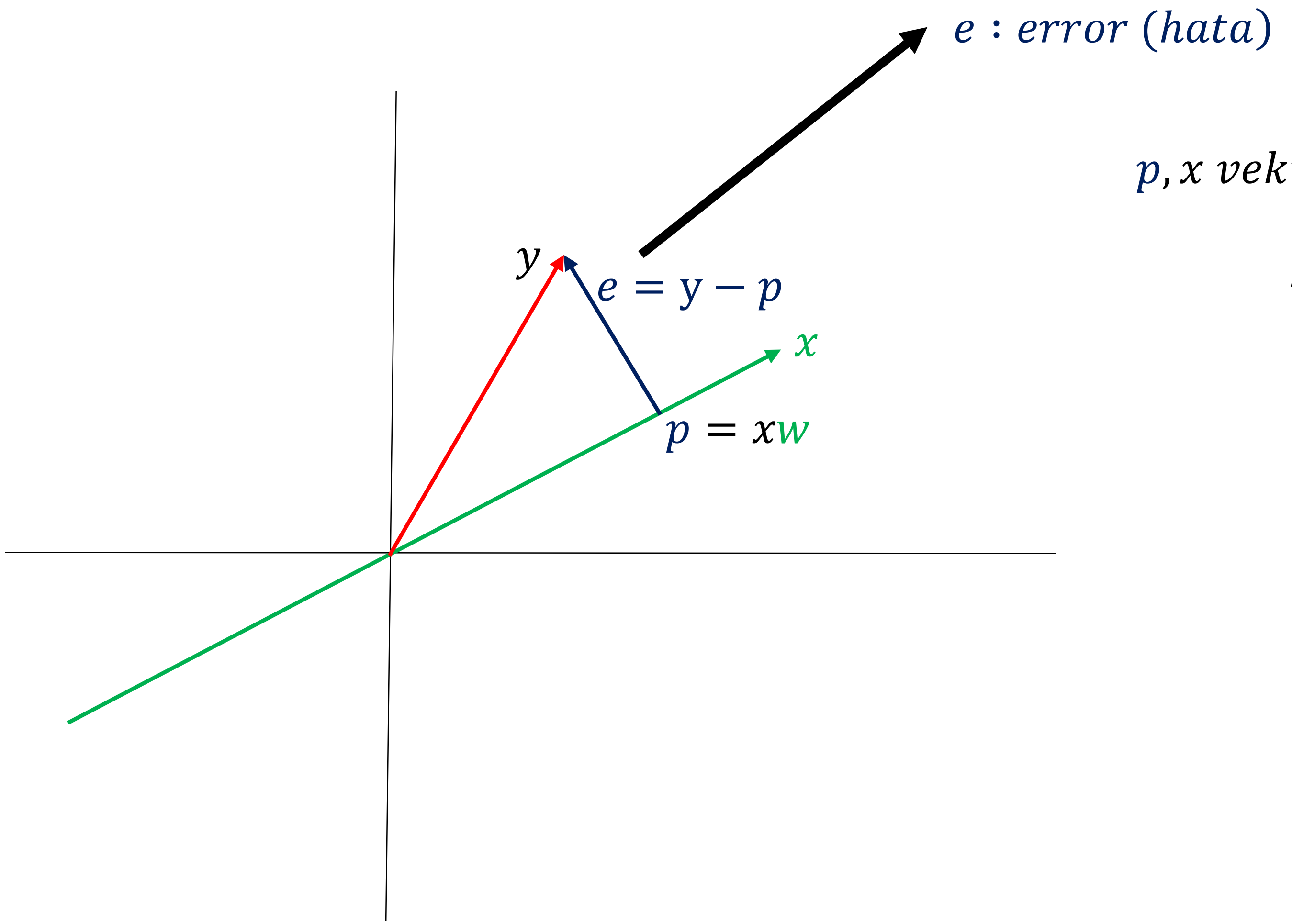


$e : error\ (hata)$

p, x *vektörünün herhangi bir katı* : $p = xw$

$x \perp e$

Projeksiyon ve Projeksiyon Matrisi

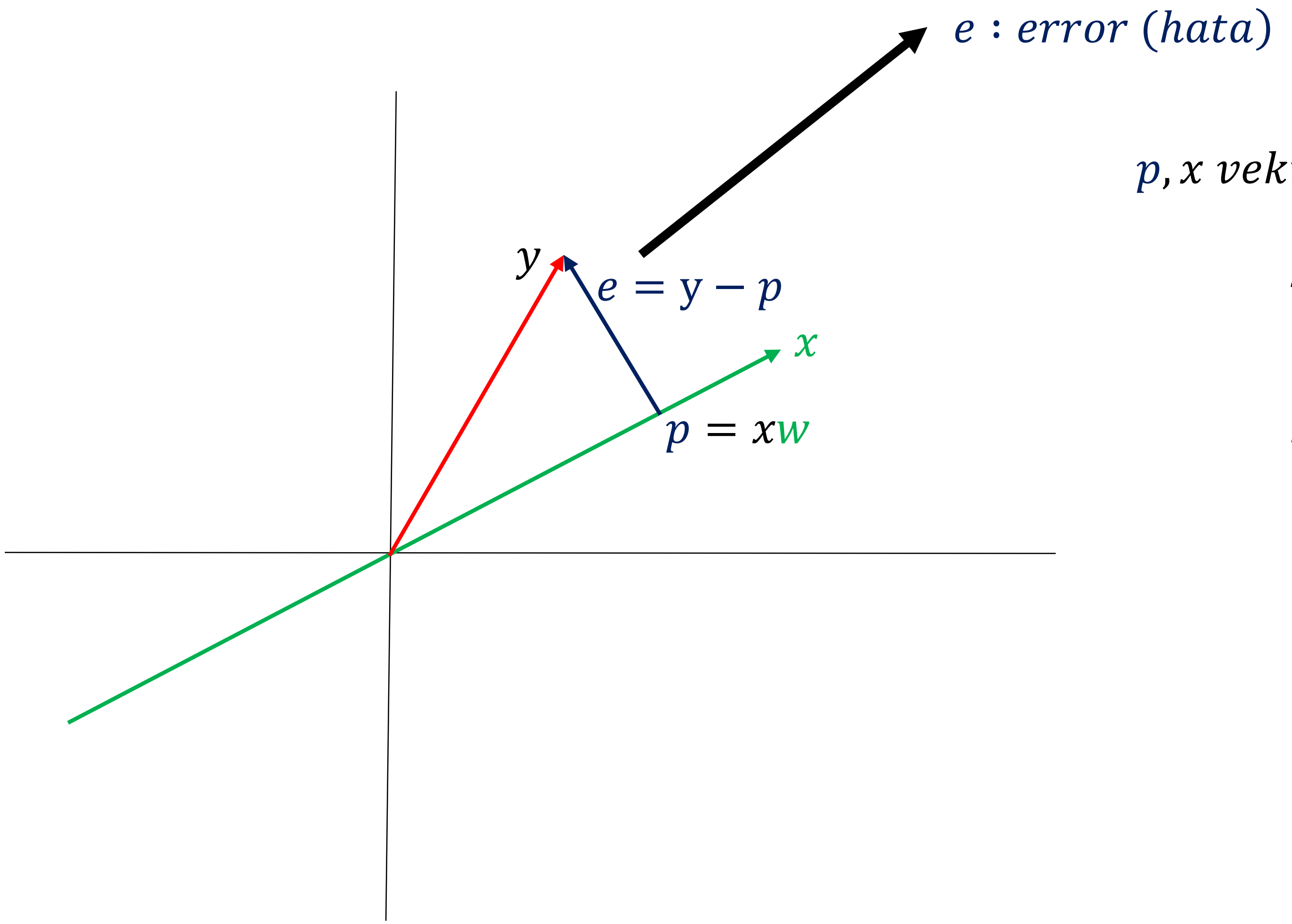


p, x *vektörünün herhangi bir katı* : $p = xw$

$x \perp e$

$x \perp (y - p)$

Projeksiyon ve Projeksiyon Matrisi



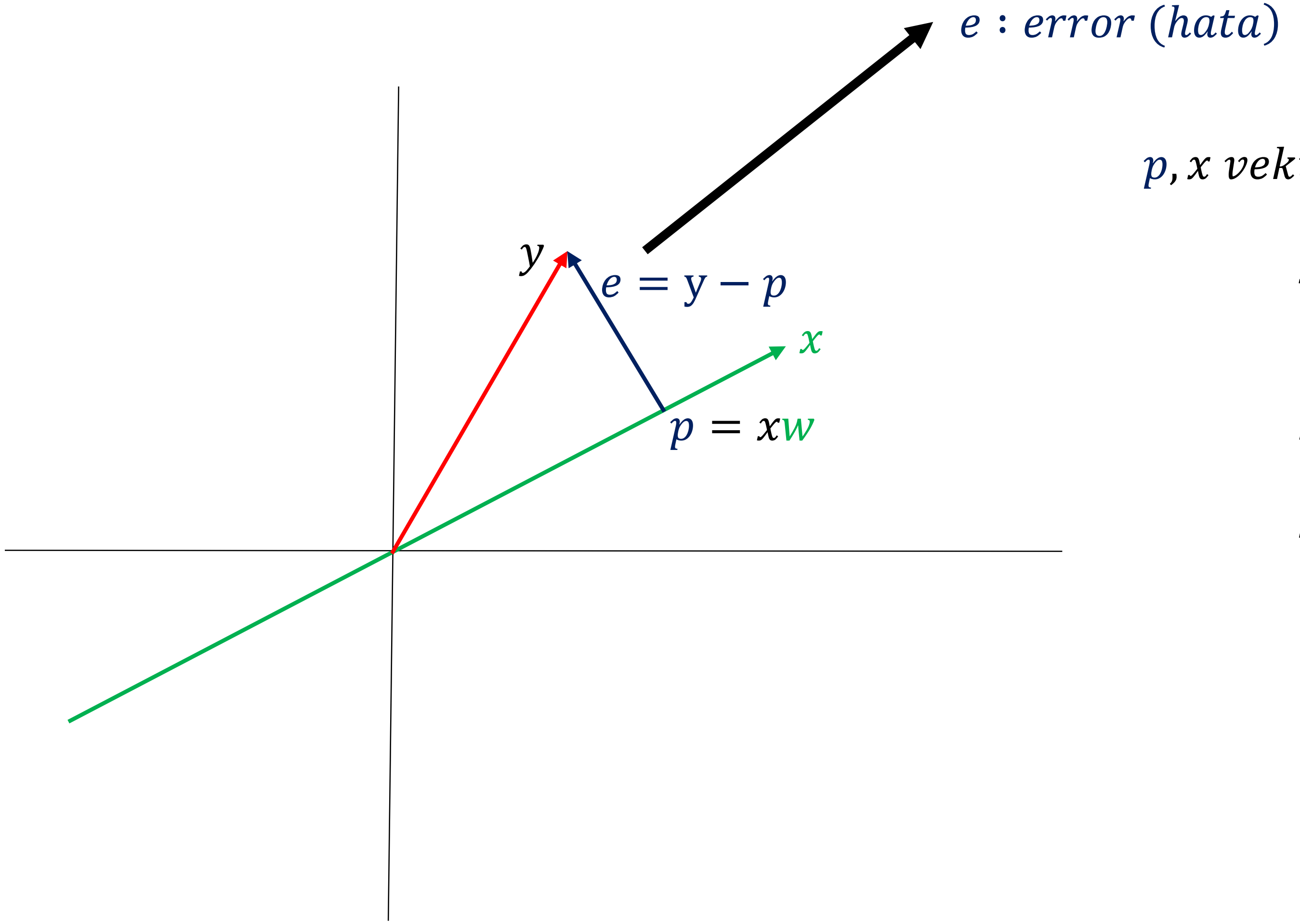
p, x *vektörünün herhangi bir katı* : $p = xw$

$$x \perp e$$

$$x \perp (y - p)$$

$$x \perp (y - xw)$$

Projeksiyon ve Projeksiyon Matrisi



p, x *vektörünün herhangi bir katı* : $p = xw$

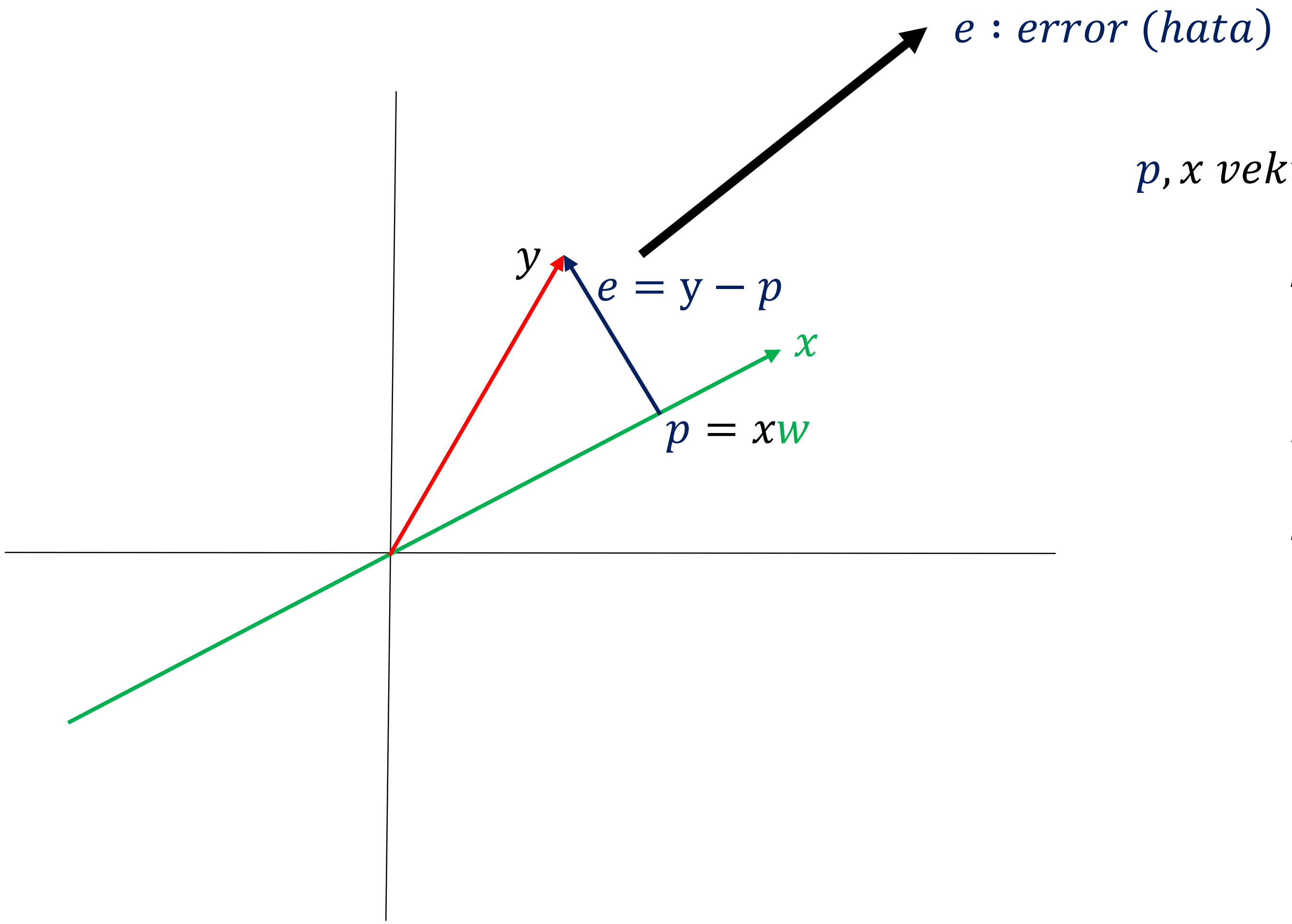
$$x \perp e$$

$$x \perp (y - p)$$

$$x \perp (y - xw)$$

$$x^T (y - xw) = 0$$

Projeksiyon ve Projeksiyon Matrisi



p, x *vektörünün herhangi bir katı* : $p = xw$

$$x \perp e$$

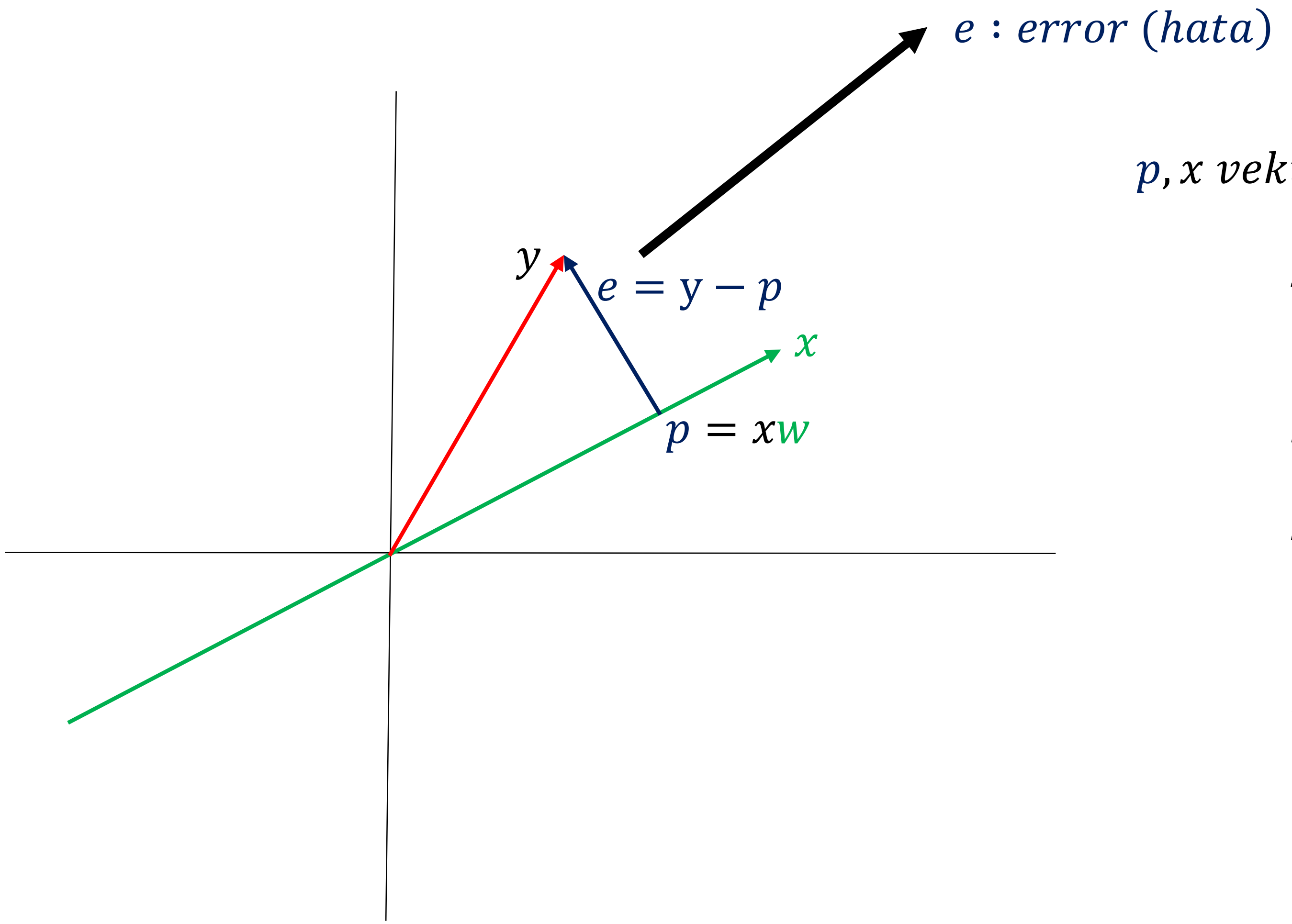
$$x \perp (y - p)$$

$$x \perp (y - xw)$$

$$x^T (y - xw) = 0$$

$$x^T y - x^T x w = 0$$

Projeksiyon ve Projeksiyon Matrisi



p, x *vektörünün herhangi bir katı* : $p = xw$

$$x \perp e$$

$$x \perp (y - p)$$

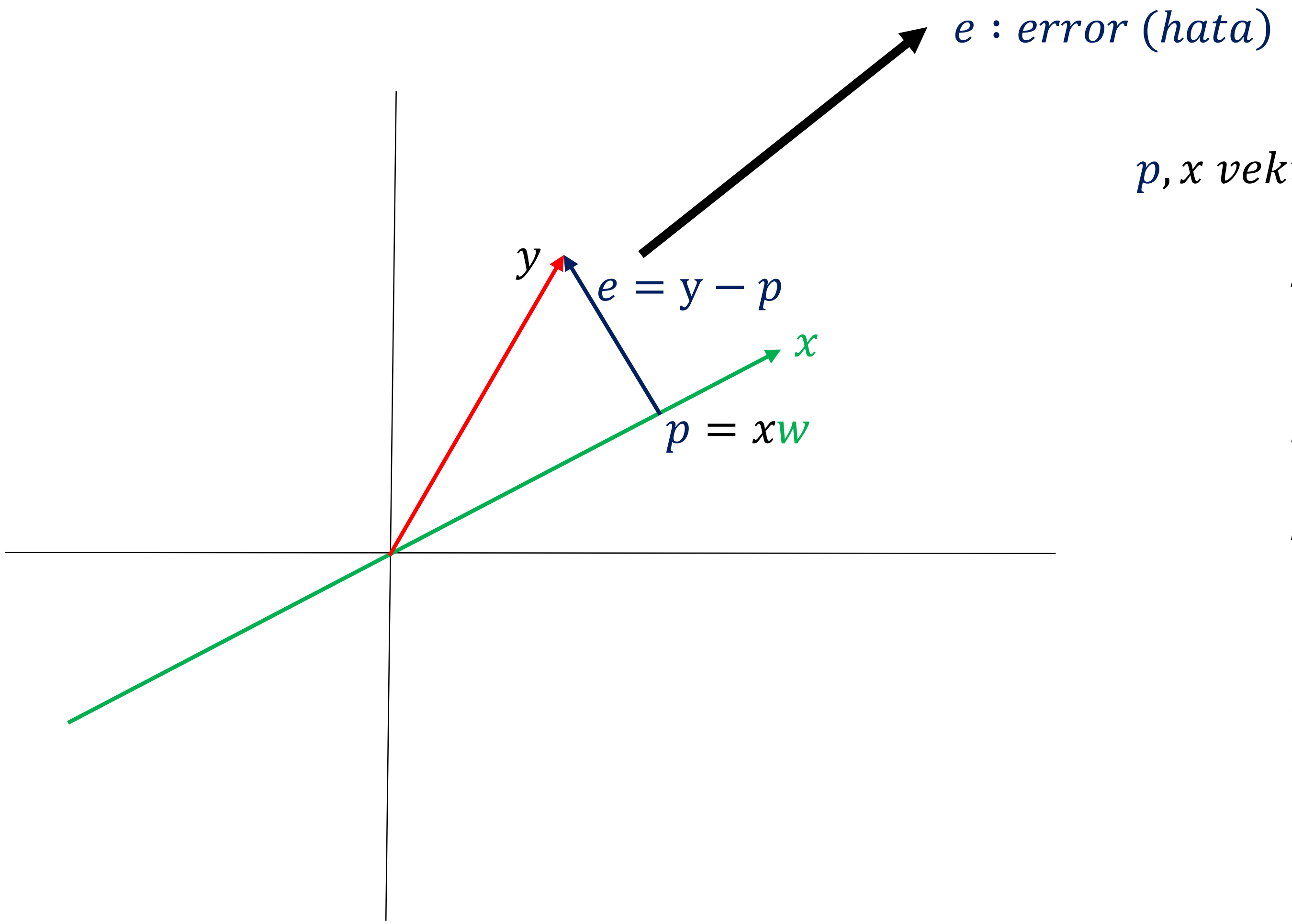
$$x \perp (y - xw)$$

$$x^T(y - xw) = 0$$

$$x^T y - x^T x w = 0$$

$$x^T x w = x^T y$$

Projeksiyon ve Projeksiyon Matrisi



p, x *vektörünün herhangi bir katı* : $p = xw$

$$x \perp e$$

$$x \perp (y - p)$$

$$x \perp (y - xw)$$

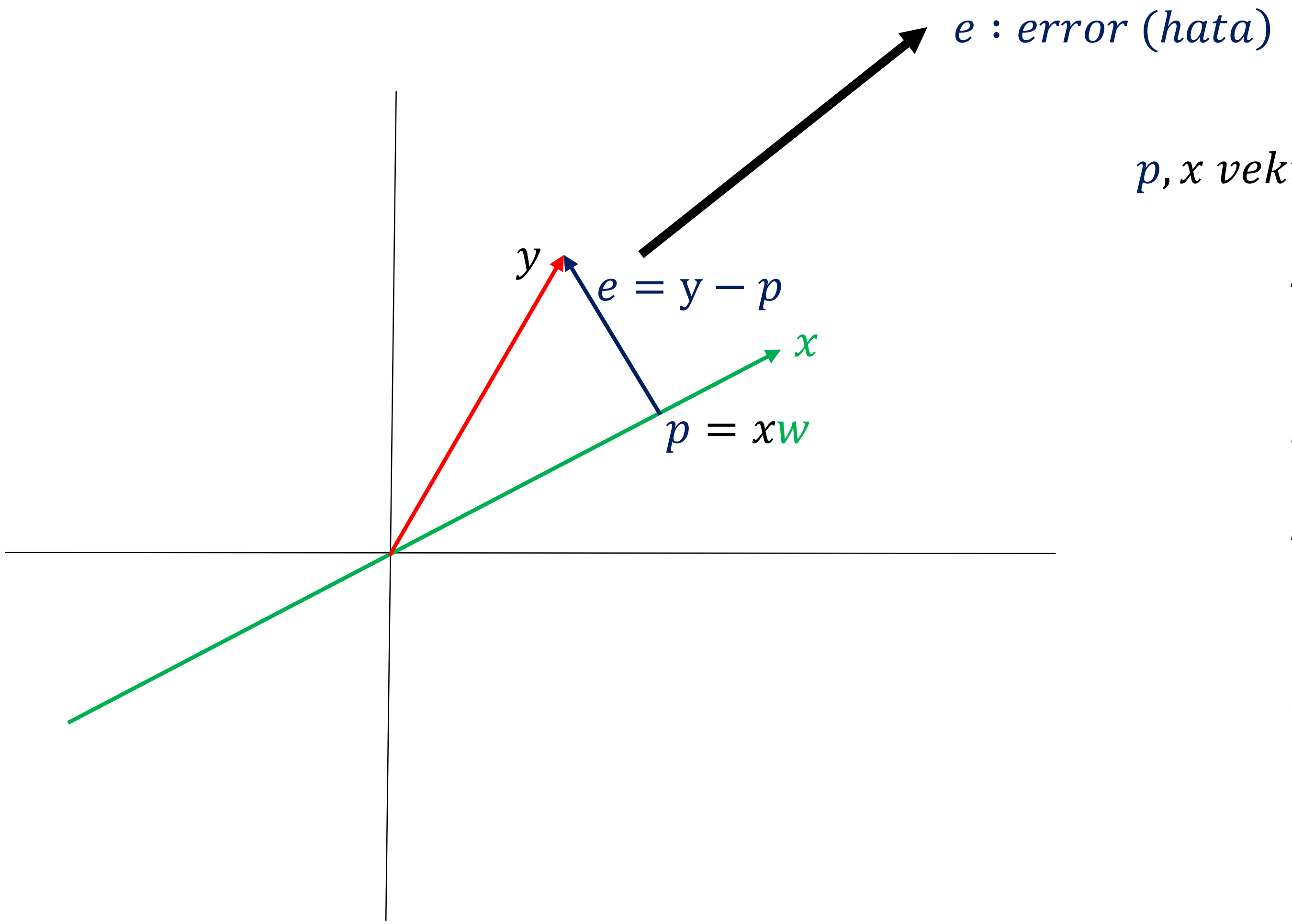
$$x^T(y - xw) = 0$$

$$x^T y - x^T xw = 0$$

$$x^T xw = x^T y$$

$$w = \frac{x^T y}{x^T x}$$

Projeksiyon ve Projeksiyon Matrisi



p, x *vektörünün herhangi bir katı* $: p = xw$

$$x \perp e$$

$$x \perp (y - p)$$

$$x \perp (y - xw)$$

$$x^T(y - xw) = 0$$

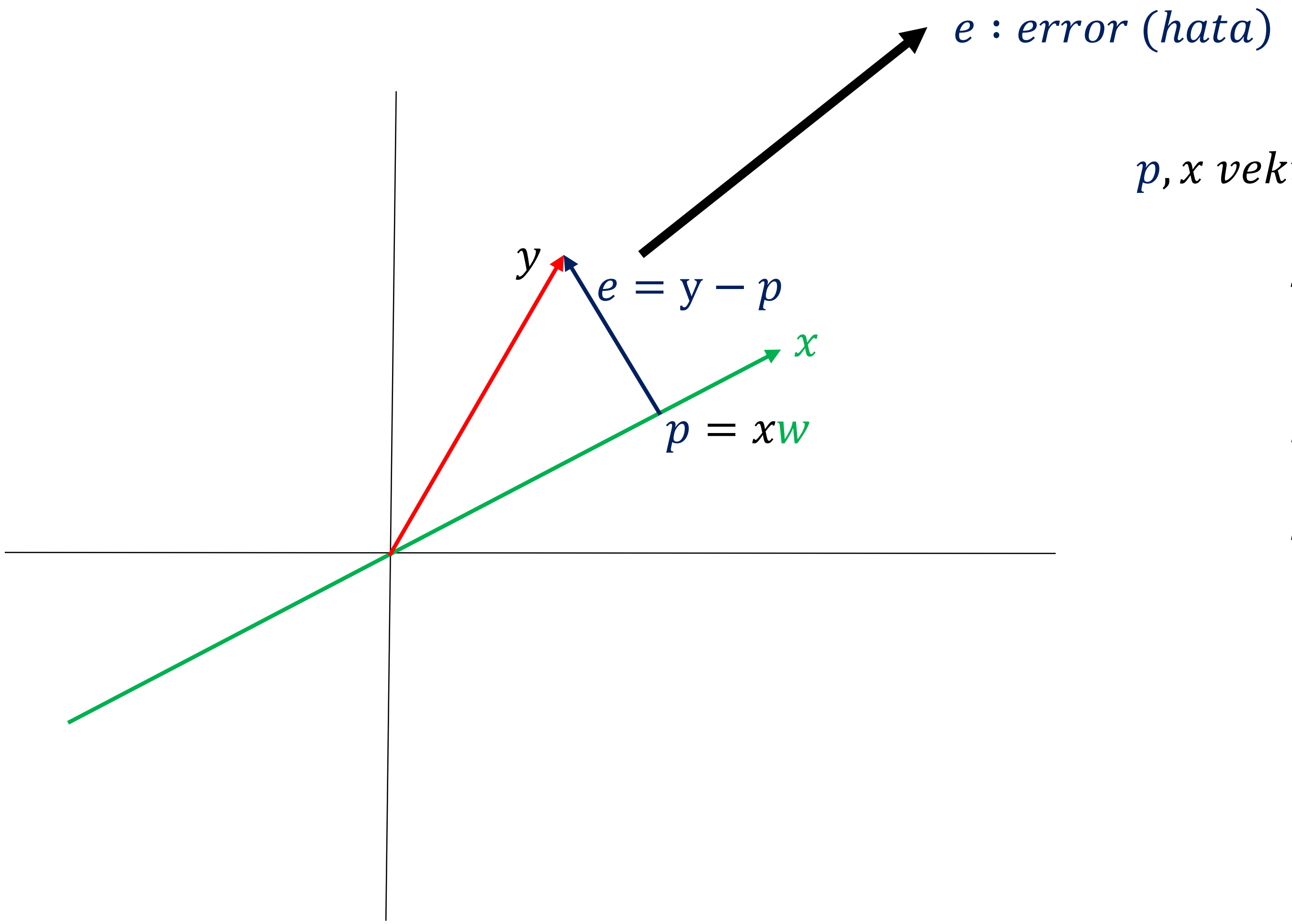
$$x^T y - x^T x w = 0$$

$$x^T x w = x^T y$$

$$w = \frac{x^T y}{x^T x}$$

$$p = xw$$

Projeksiyon ve Projeksiyon Matrisi



p, x *vektörünün herhangi bir katı* : $p = xw$

$$x \perp e$$

$$x \perp (y - p)$$

$$x \perp (y - xw)$$

$$x^T(y - xw) = 0$$

$$x^T y - x^T x w = 0$$

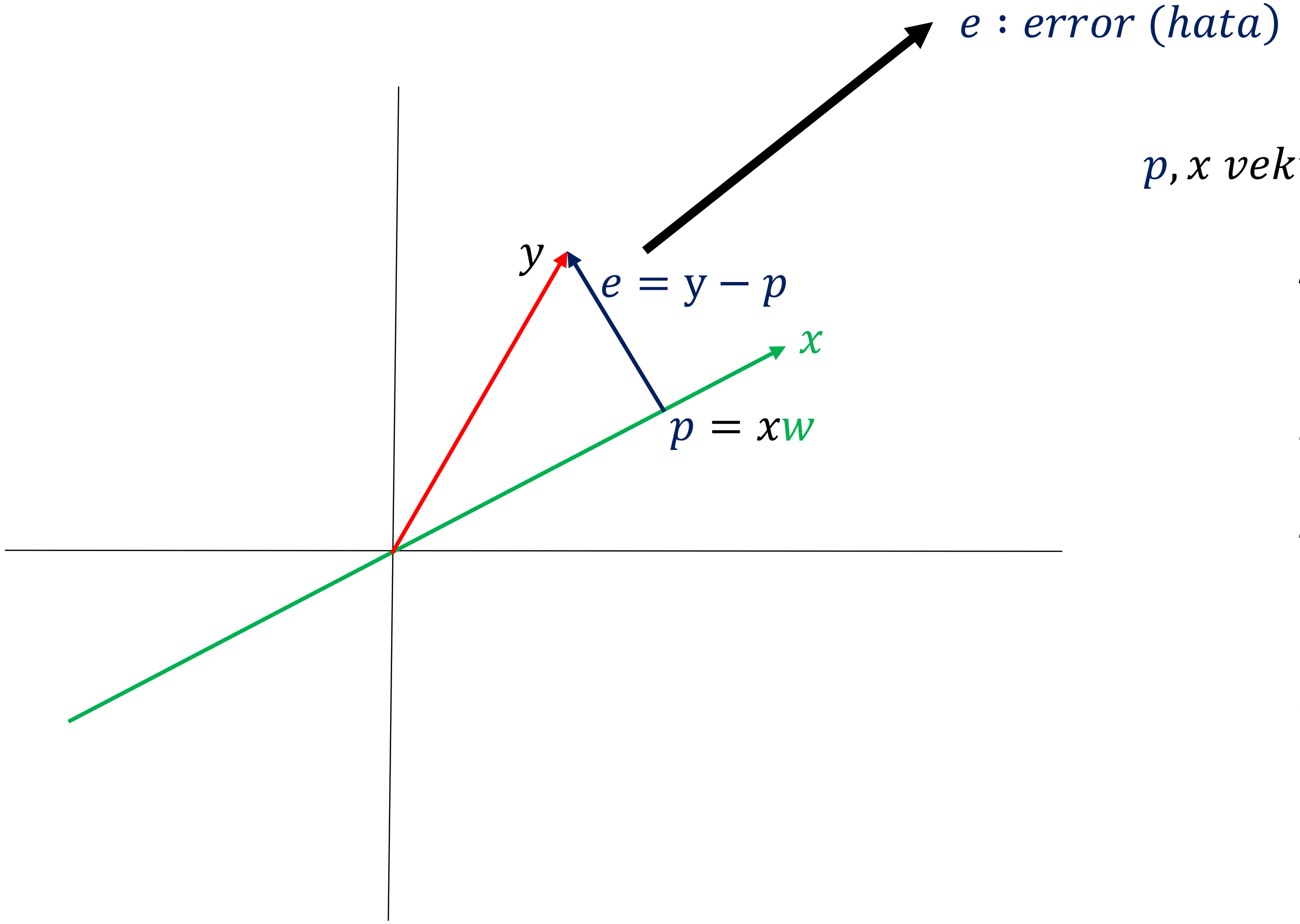
$$x^T x w = x^T y$$

$$w = \frac{x^T y}{x^T x}$$

projeksiyon

$$p = xw \rightarrow \boxed{p = x\frac{x^T y}{x^T x} = Py}$$

Projeksiyon ve Projeksiyon Matrisi



p, x *vektörünün herhangi bir katı* : $p = xw$

$$x \perp e$$

$$x \perp (y - p)$$

$$x \perp (y - xw)$$

$$x^T (y - xw) = 0$$

$$x^T y - x^T x w = 0$$

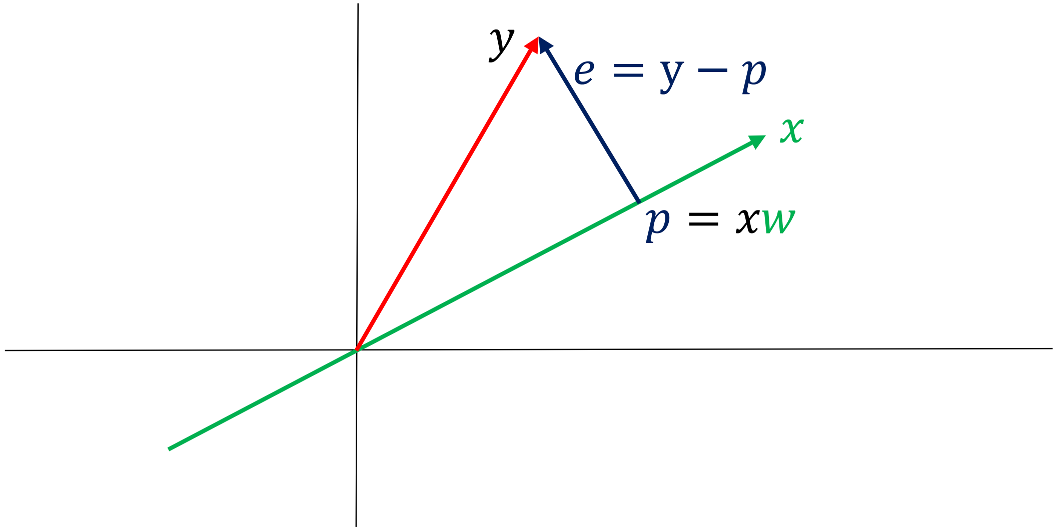
$$x^T x w = x^T y$$

$$w = \frac{x^T y}{x^T x}$$

projeksiyon — *projeksiyon matrisi*

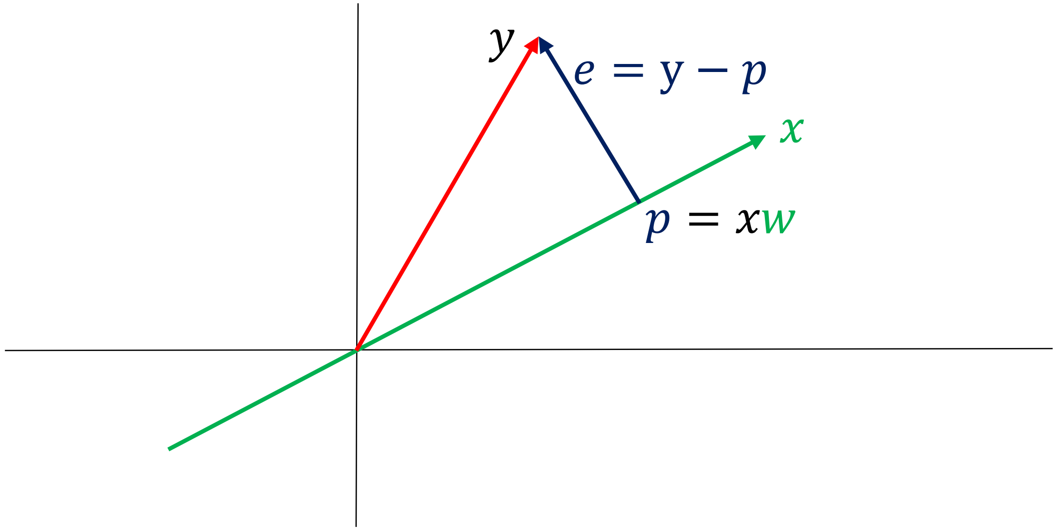
$$p = xw \rightarrow \boxed{p = x\frac{x^T y}{x^T x} = Py} \rightarrow \boxed{P = \frac{xx^T}{x^T x}}$$

Projeksiyon ve Projeksiyon Matrisi



$\ddot{O}nemli\ \ddot{O}zellikler;$

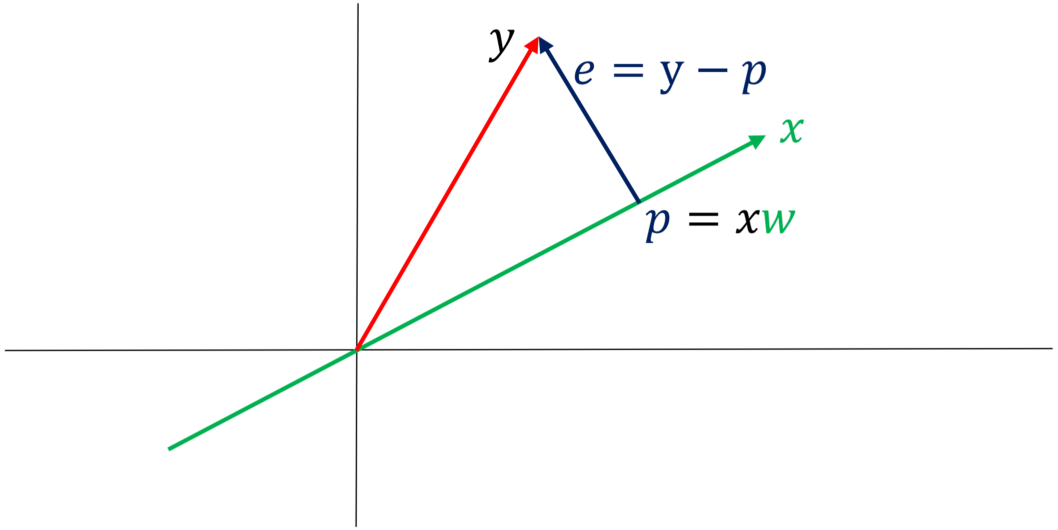
Projeksiyon ve Projeksiyon Matrisi



$Önemli\ Özellikler;$

1) $Projeksiyon\ matrisi\ simekriktir.$

Projeksiyon ve Projeksiyon Matrisi

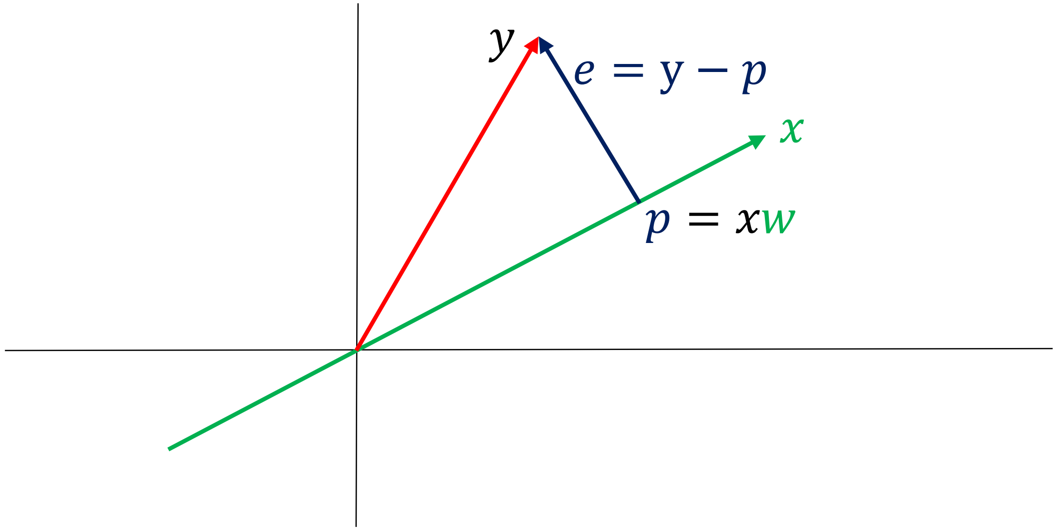


$Önemli\ Özellikler;$

$1)\ Projeksiyon\ matrisi\ simekriktir.$

$$P^T = \frac{xx^{T^T}}{x^T x} = \frac{x^{TT} x^T}{x^T x^{TT}} = \frac{xx^T}{x^T x} = P$$

Projeksiyon ve Projeksiyon Matrisi



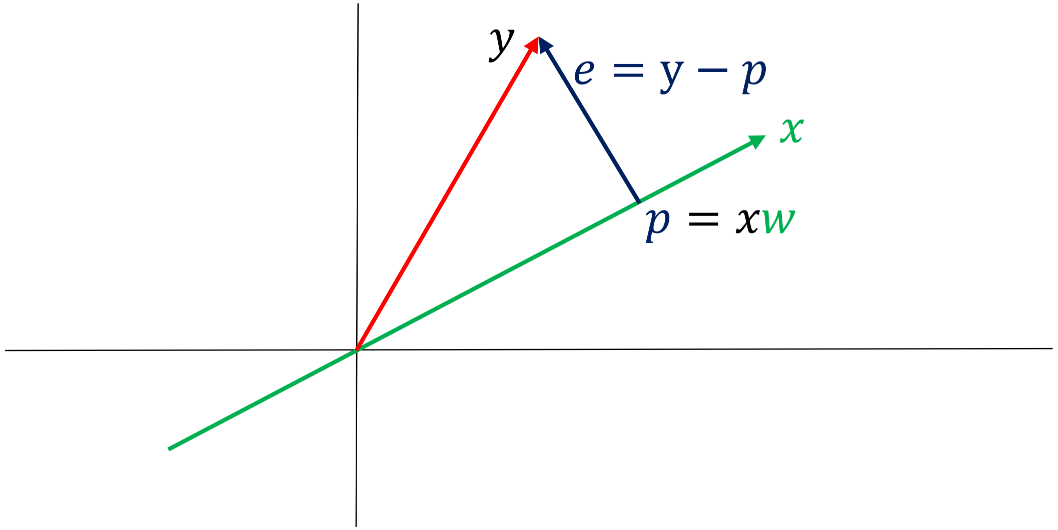
Önemli Özellikler;

1) *Projeksiyon matrisi simekriktir.*

$$P^T = \frac{{xx^T}^T}{x^T x} = \frac{x^{TT} x^T}{x^T x^{TT}} = \frac{xx^T}{x^T x} = P$$

2) *Eğer iki kez projeksiyon yaparsam, aynı yerde kalmış olurum.*

Projeksiyon ve Projeksiyon Matrisi



$Önemli\ Özellikler;$

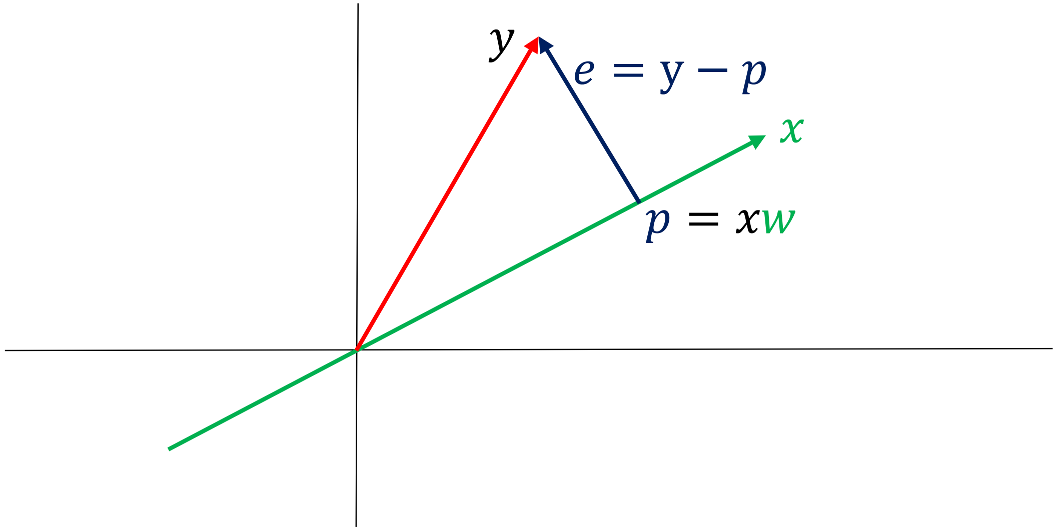
$1)\ Projeksiyon\ matrisi\ simekriktir.$

$$P^T = \frac{{xx^T}^T}{x^Tx} = \frac{x^{TT}x^T}{x^Tx^{TT}} = \frac{xx^T}{x^Tx} = P$$

$2)\ Eğer\ iki\ kez\ projeksiyon\ yaparsam,$
$aynı\ yerde\ kalmış\ olurum.$

$$P^2 = \frac{xx^Txx^T}{x^Txx^Tx} = \frac{(x)(x^Tx)(x^T)}{(x^Tx)(x^Tx)} = \frac{xx^T}{x^Tx} = P$$

Projeksiyon ve Projeksiyon Matrisi



$$P = \frac{xx^T}{x^Tx}$$

Önemli Özellikler;

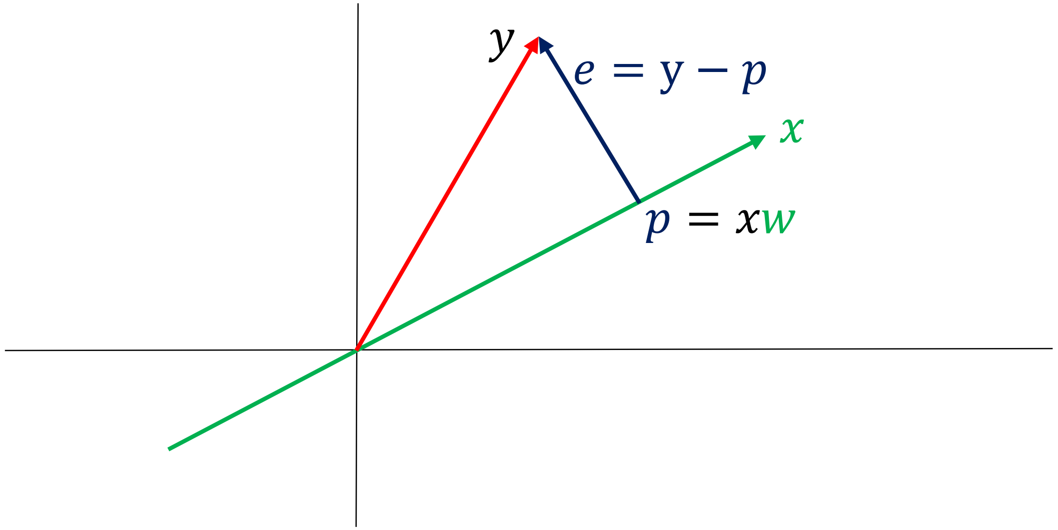
1) *Projeksiyon matrisi simekriktir.*

$$P^T = \frac{{xx^T}^T}{x^Tx} = \frac{x^{TT}x^T}{x^Tx^{TT}} = \frac{xx^T}{x^Tx} = P$$

2) *Eğer iki kez projeksiyon yaparsam, aynı yerde kalmış olurum.*

$$P^2 = \frac{xx^Txx^T}{x^Txx^Tx} = \frac{(x)(x^Tx)(x^T)}{(x^Tx)(x^Tx)} = \frac{xx^T}{x^Tx} = P$$

Projeksiyon ve Projeksiyon Matrisi



$$P = \frac{xx^T}{x^Tx}$$ → *matris*

Önemli Özellikler;

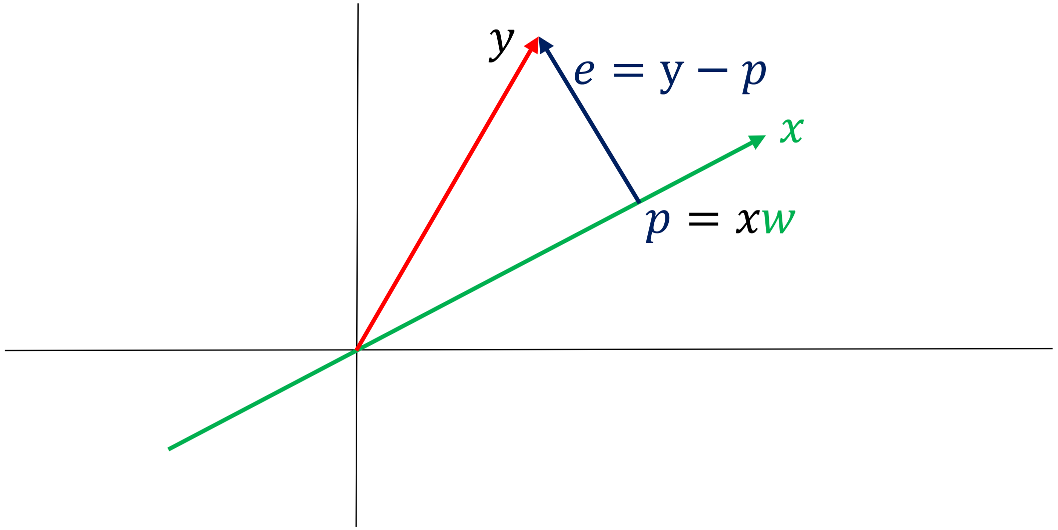
1) *Projeksiyon matrisi simekriktir.*

$$P^T = \frac{{xx^T}^T}{x^Tx} = \frac{x^{TT}x^T}{x^Tx^{TT}} = \frac{xx^T}{x^Tx} = P$$

2) *Eğer iki kez projeksiyon yaparsam, aynı yerde kalmış olurum.*

$$P^2 = \frac{xx^Txx^T}{x^Txx^Tx} = \frac{(x)(x^Tx)(x^T)}{(x^Tx)(x^Tx)} = \frac{xx^T}{x^Tx} = P$$

Projeksiyon ve Projeksiyon Matrisi



$$P = \frac{xx^T}{x^Tx}$$

$xx^T \rightarrow$ *matris*

$x^Tx \rightarrow$ *dot product : iç çarpım* $(\|x\|^2)$

Önemli Özellikler;

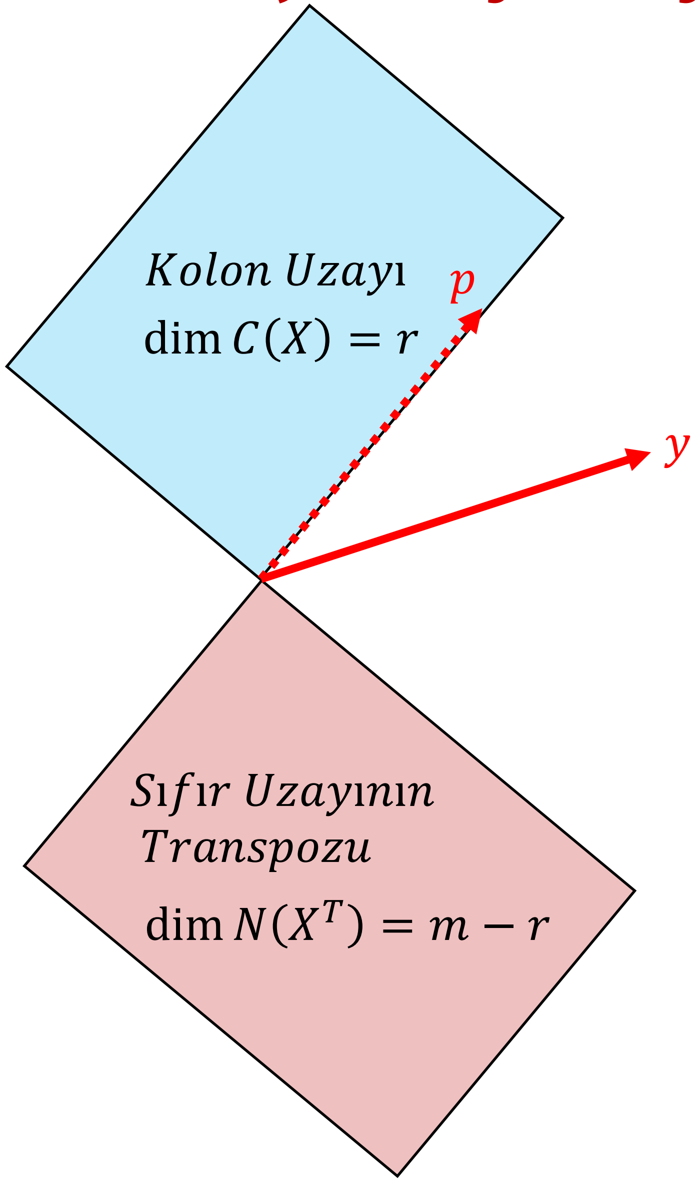
1) *Projeksiyon matrisi simekriktir.*

$$P^T = \frac{{xx^T}^T}{x^Tx} = \frac{x^{TT}x^T}{x^Tx^{TT}} = \frac{xx^T}{x^Tx} = P$$

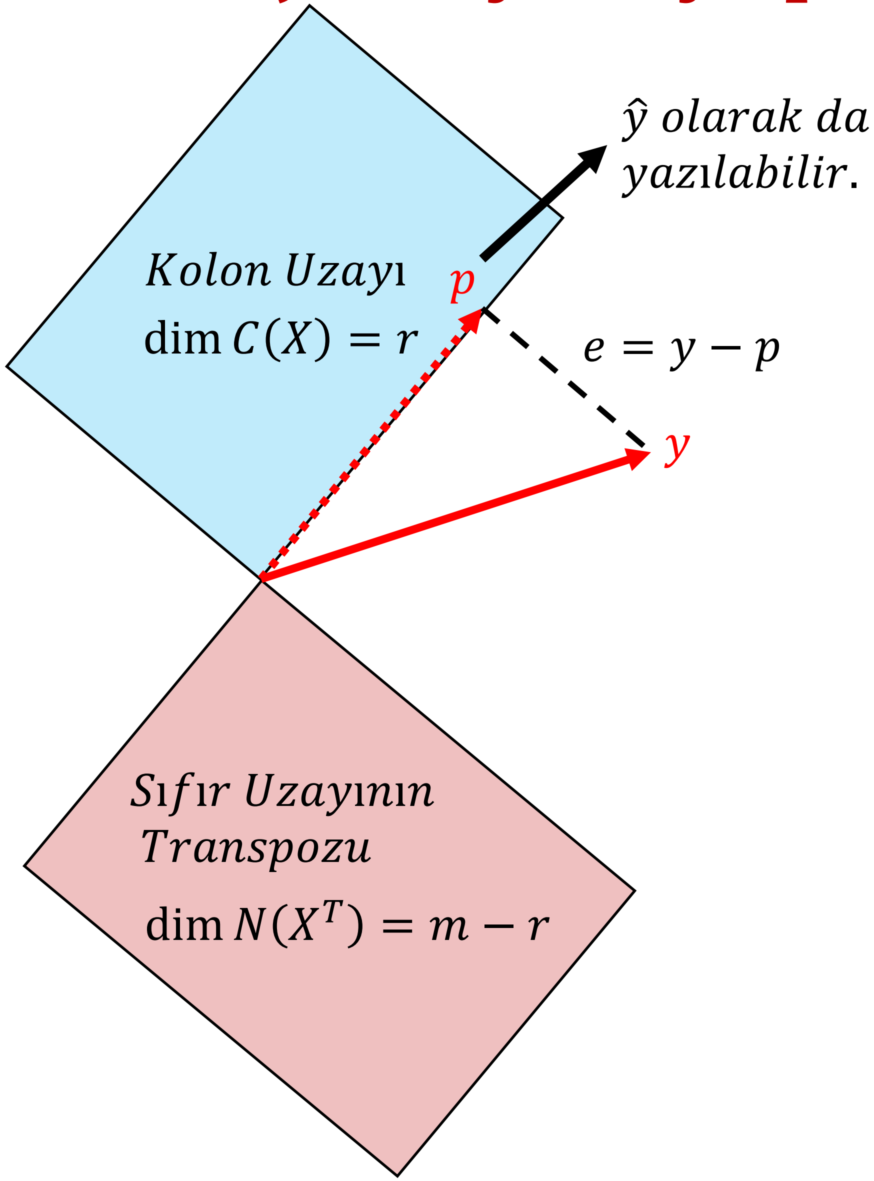
2) *Eğer iki kez projeksiyon yaparsam, aynı yerde kalmış olurum.*

$$P^2 = \frac{xx^Txx^T}{x^Txx^Tx} = \frac{(x)(x^Tx)(x^T)}{(x^Tx)(x^Tx)} = \frac{xx^T}{x^Tx} = P$$

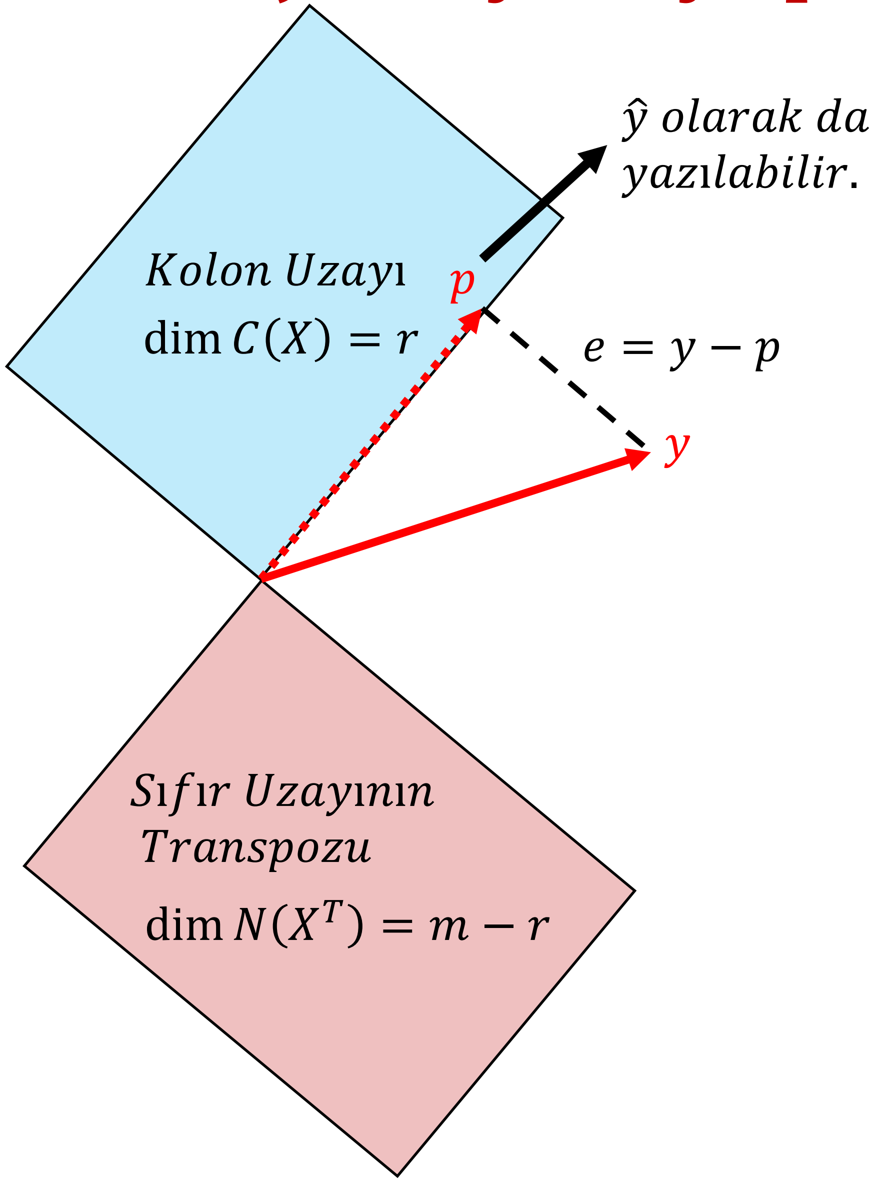
Neden Projeksiyon yaparız ?



Neden Projeksiyon yaparız ?

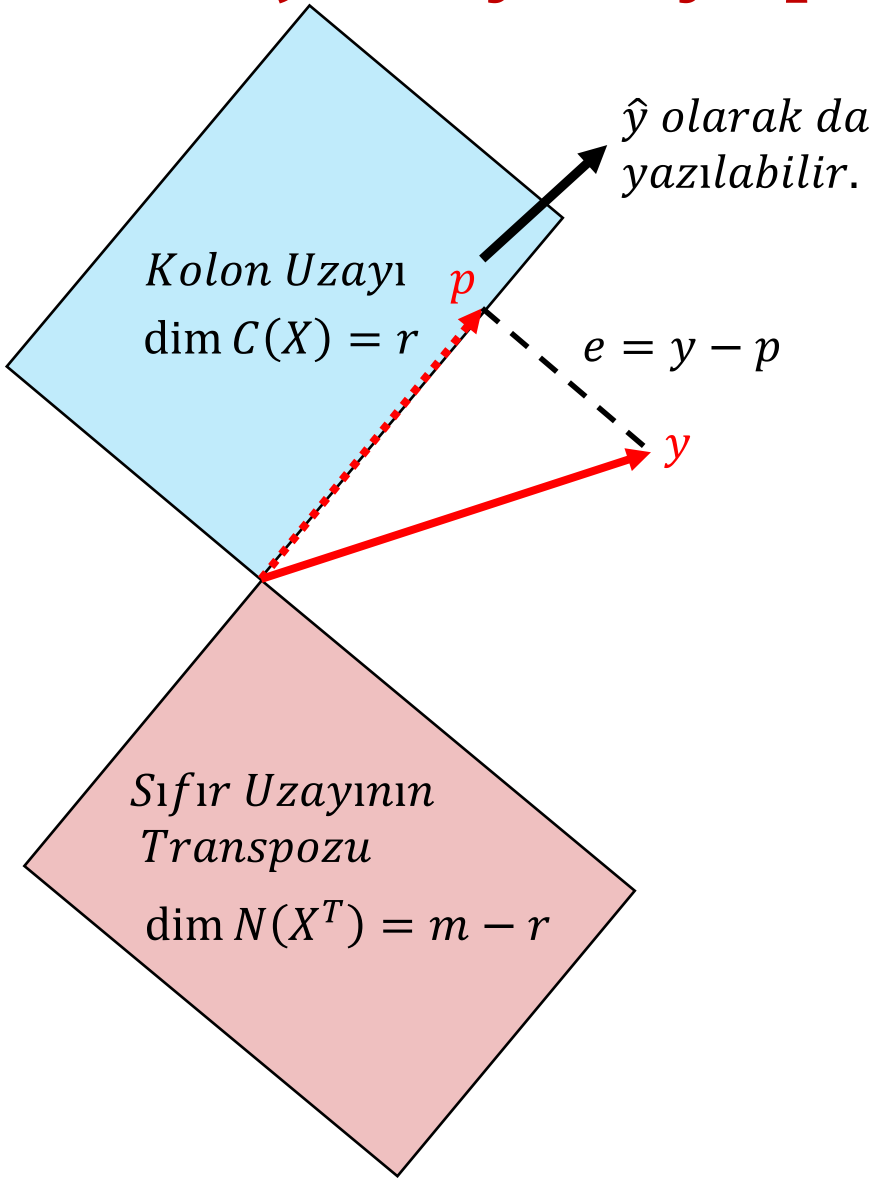


Neden Projeksiyon yaparız ?



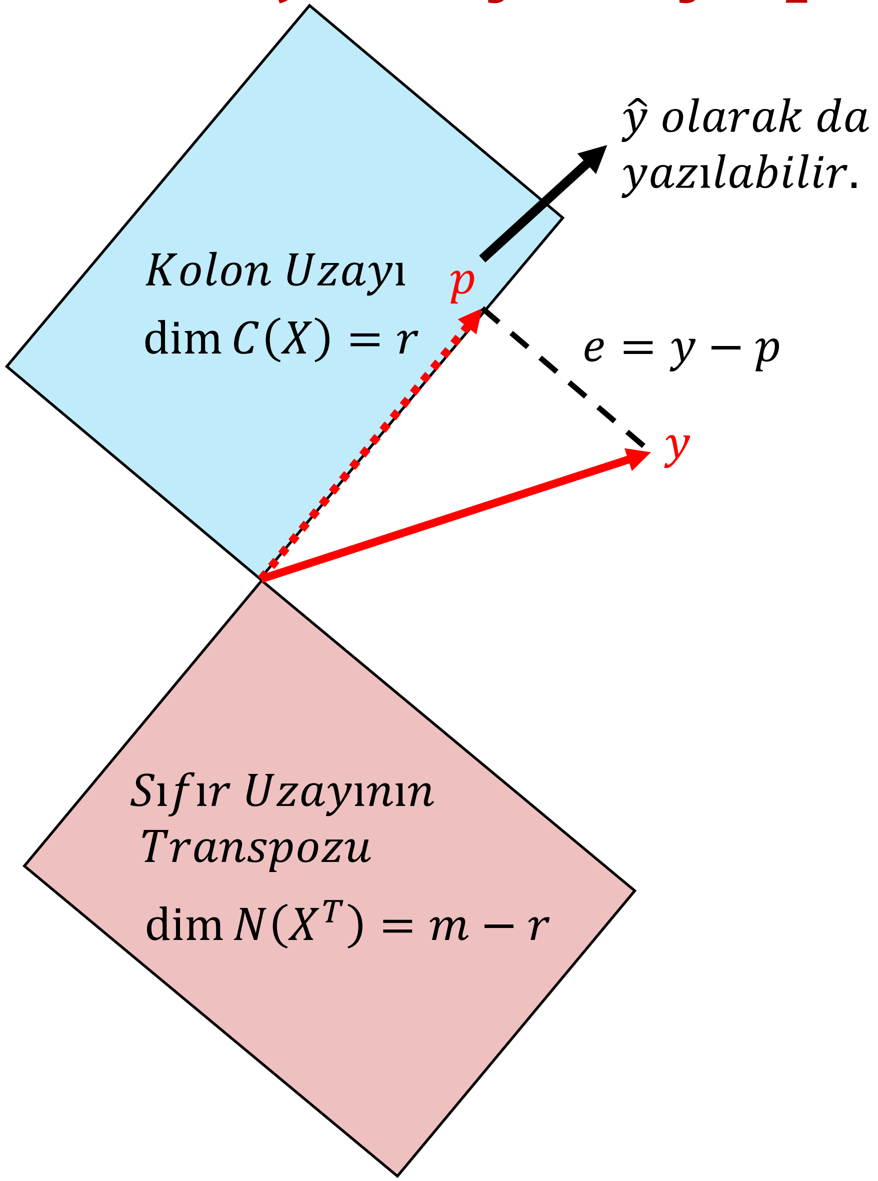
- Çünkü $Xw = y$ denkleminin çözümü olmayabilir.

Neden Projeksiyon yaparız ?



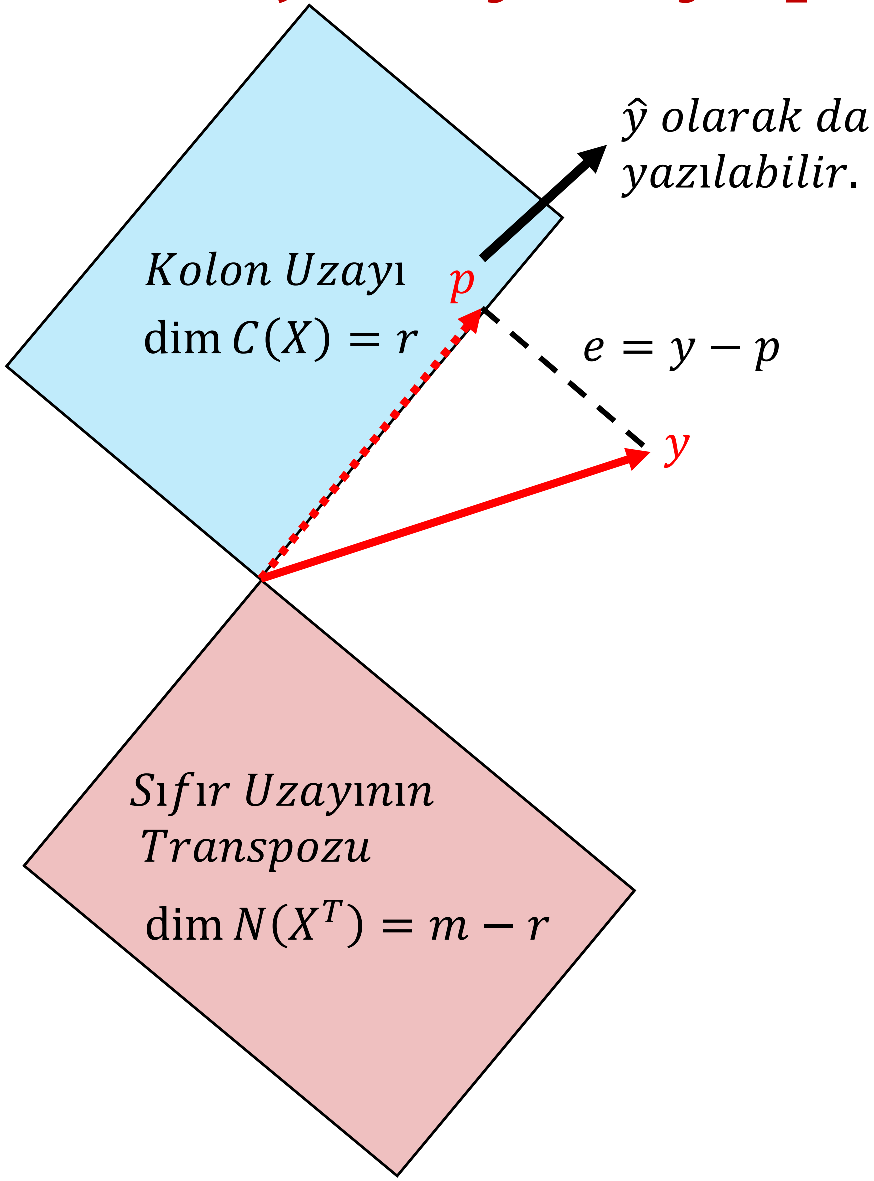
- Çünkü $Xw = y$ *denkleminin* çözümü *olmayabilir*.
- *yani uygun y değerlerini alamayabilirim*.

Neden Projeksiyon yaparız ?



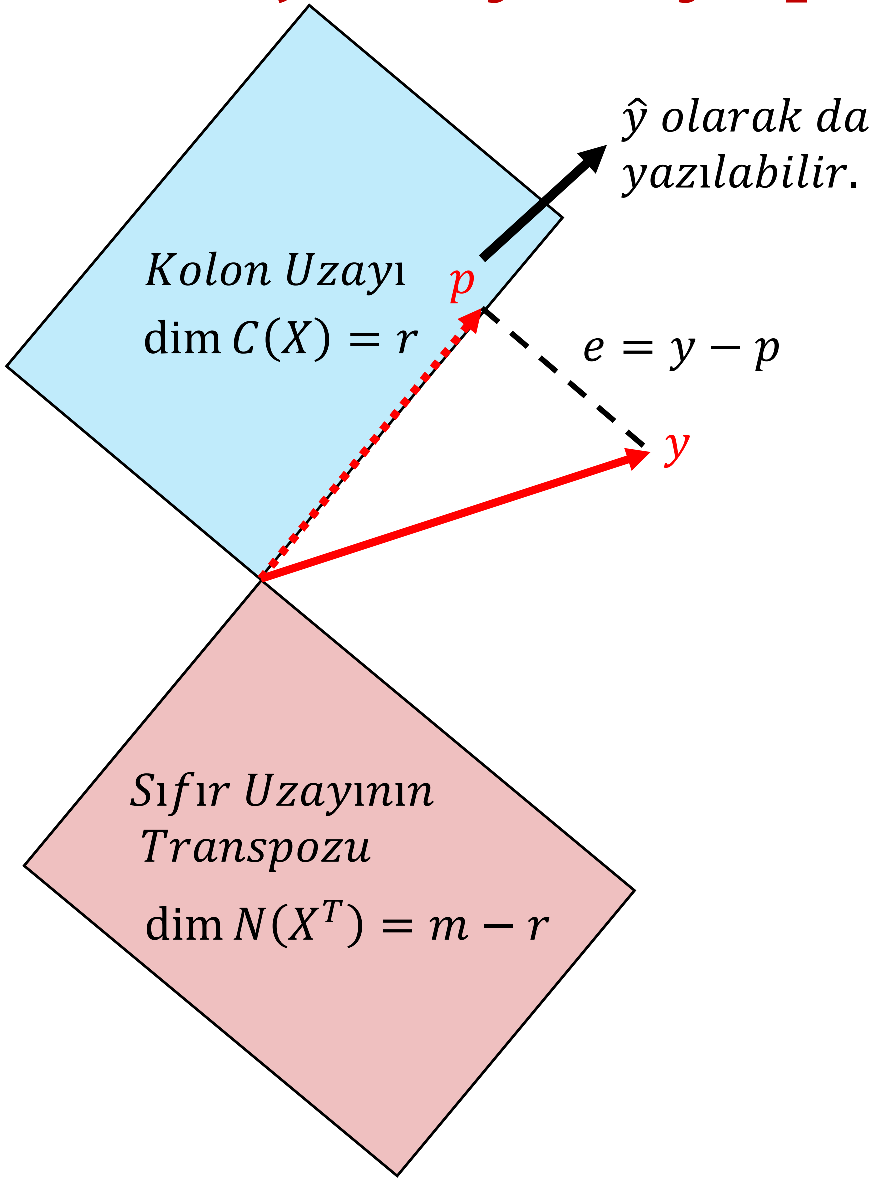
- Çünkü $Xw = y$ denkleminin çözümü olmayabilir.
- yani uygun y değerlerini alamayabilirim.
- Kolon uzayında çalışmam gerekli. Fakat kolon uzayında değilim.

Neden Projeksiyon yaparız ?



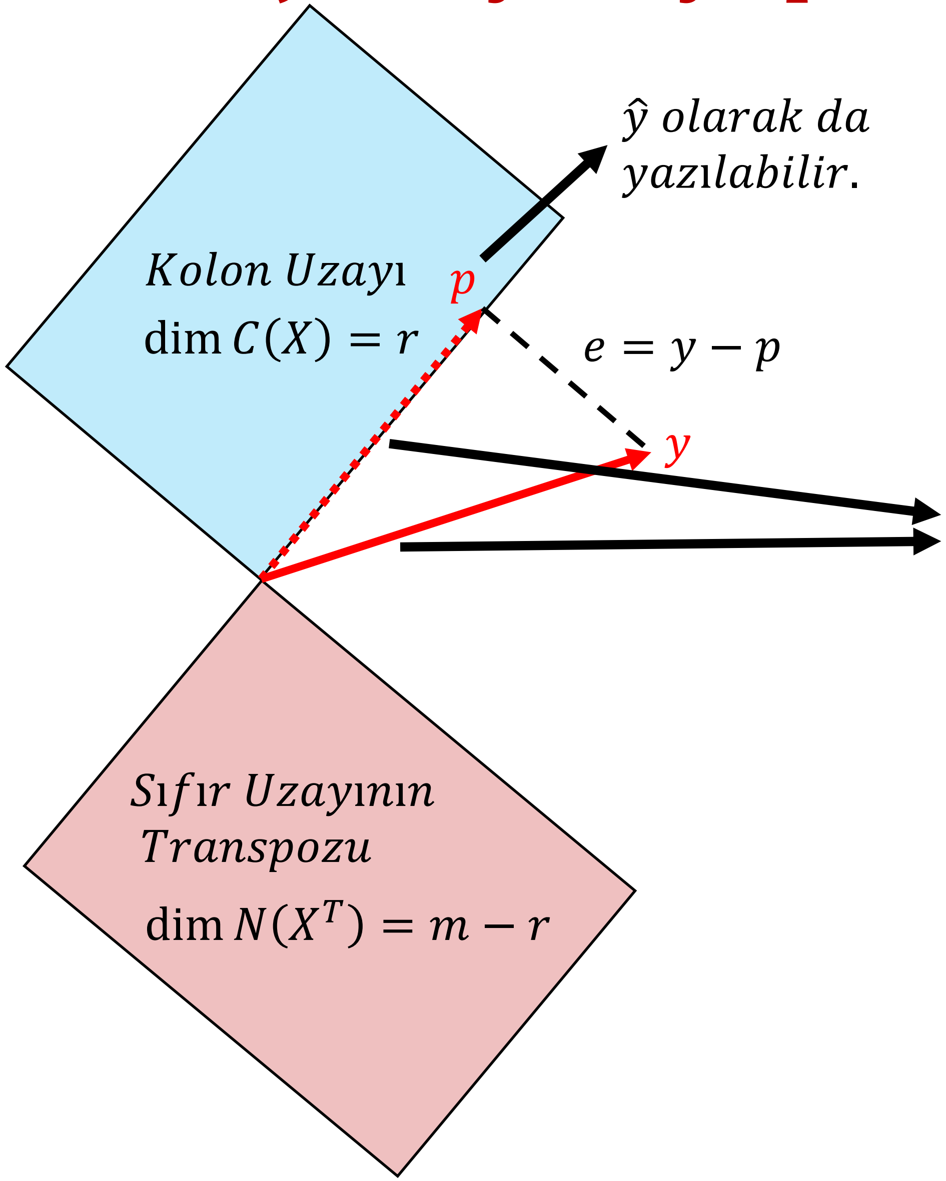
- Çünkü $Xw = y$ denkleminin çözümü olmayabilir.
- yani uygun y değerlerini alamayabilirim.
- Kolon uzayında çalışmam gerekli. Fakat kolon uzayında değilim.
- Sıfır uzayında da değilim. İkisinin arasında bir yerdeyim.

Neden Projeksiyon yaparız ?



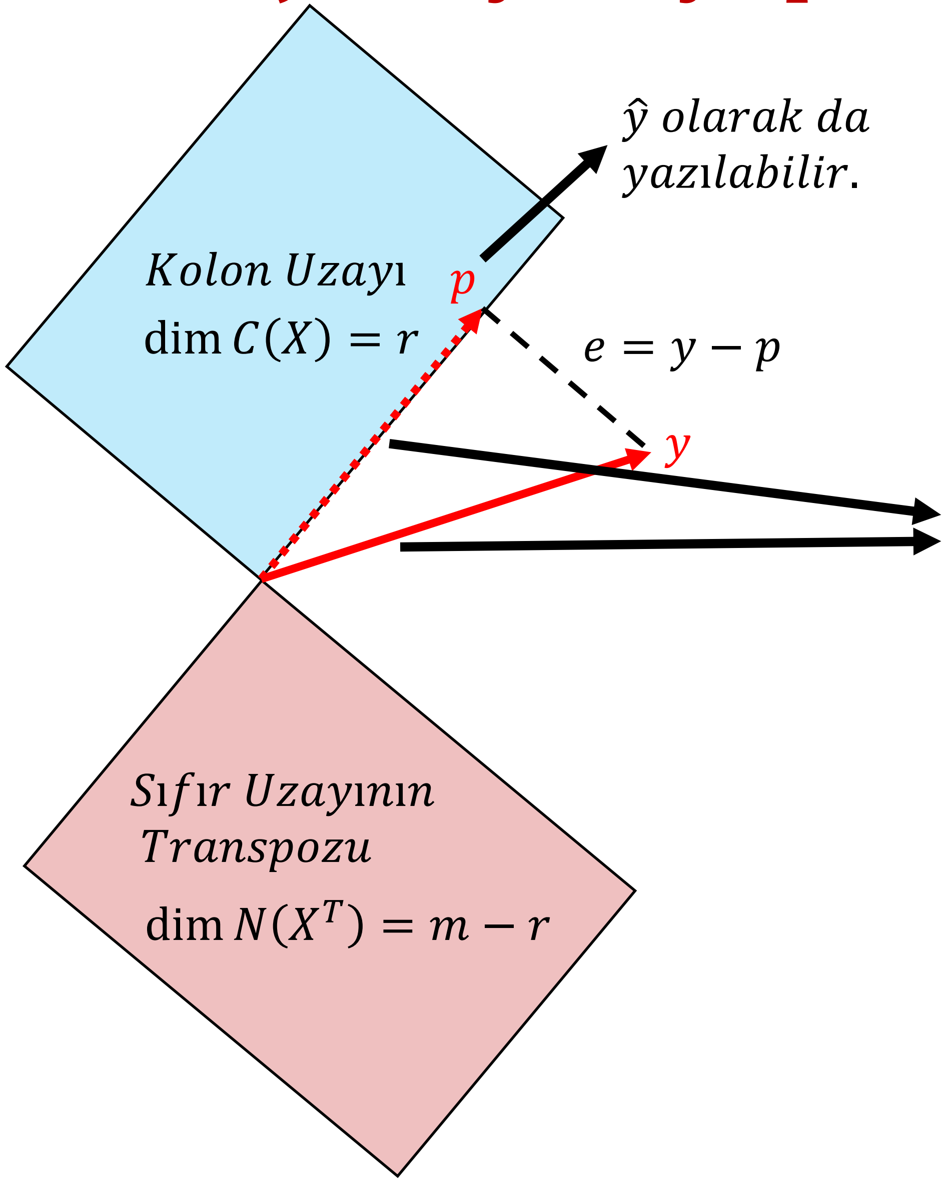
- Çünkü $Xw = y$ *denkleminin* çözümü *olmayabilir*.
- *yani uygun y değerlerini alamayabilirim.*
- *Kolon uzayında çalışmam gerekli.*
  Fakat kolon uzayında değilim.
- *Sıfır uzayında da değilim.*
  İkisinin arasında bir yerdeyim.
- *Dolayısı ile kolon uzayında çalışmam gerektiğinden,*
  kolon uzayına projeksiyon yapmalıyım.

Neden Projeksiyon yaparız ?



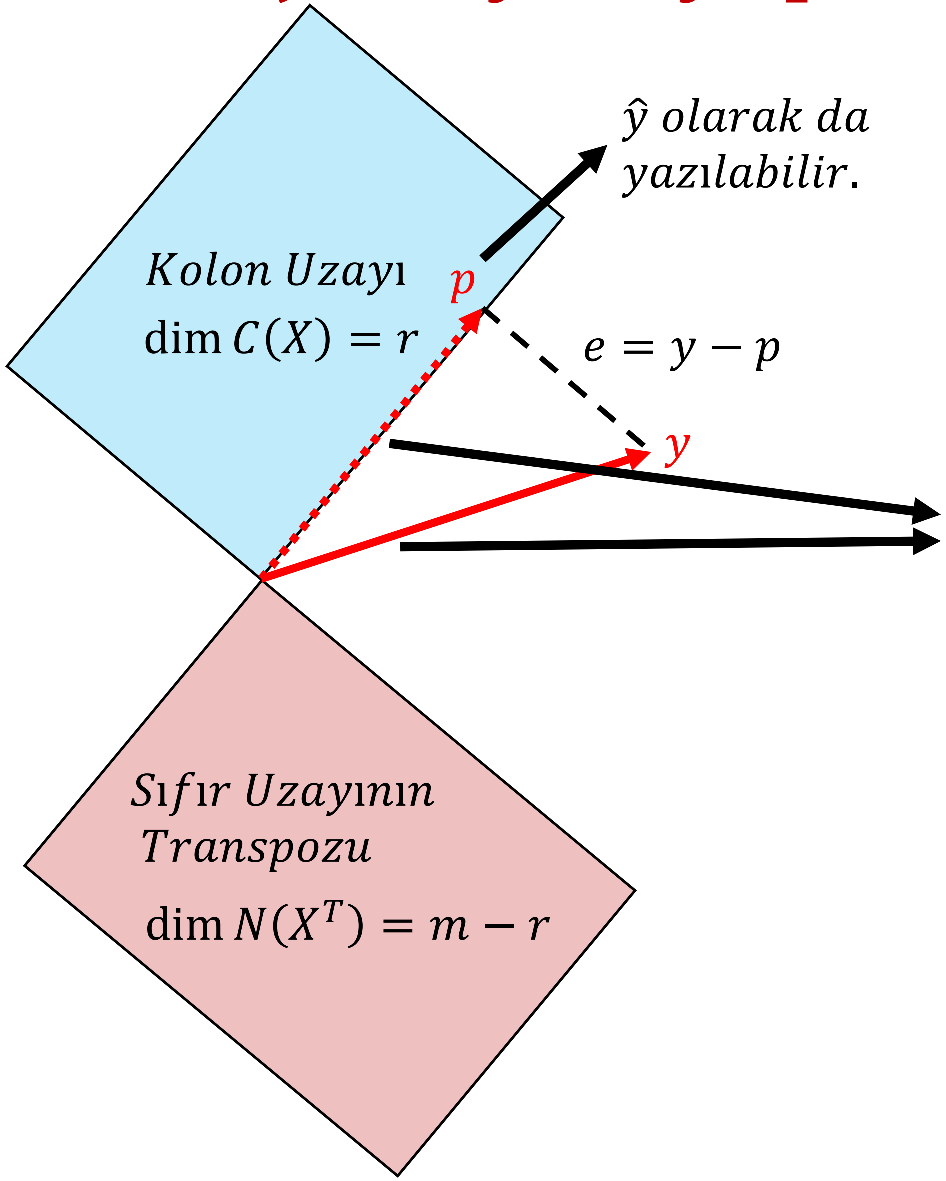
- Çünkü $Xw = y$ denkleminin çözümü olmayabilir.
- yani uygun y değerlerini alamayabilirim.
- Kolon uzayında çalışmam gerekli.
  Fakat kolon uzayında değilim.
- Sıfır uzayında da değilim.
  İkisinin arasında bir yerdeyim.
- Dolayısı ile kolon uzayında çalışmam gerektiğinden,
  kolon uzayına projeksiyon yapmalıyım.

Neden Projeksiyon yaparız ?



- Çünkü $Xw = y$ denkleminin çözümü olmayabilir.
- yani uygun y değerlerini alamayabilirim.
- Kolon uzayında çalışmam gerekli. Fakat kolon uzayında değilim.
- Sıfır uzayında da değilim. İkisinin arasında bir yerdeyim.
- Dolayısı ile kolon uzayında çalışmam gerektiğinden, kolon uzayına projeksiyon yapmalıyım.
- Bundan dolayı bizde; $X\hat{w} = p$ denklemini çözeriz.

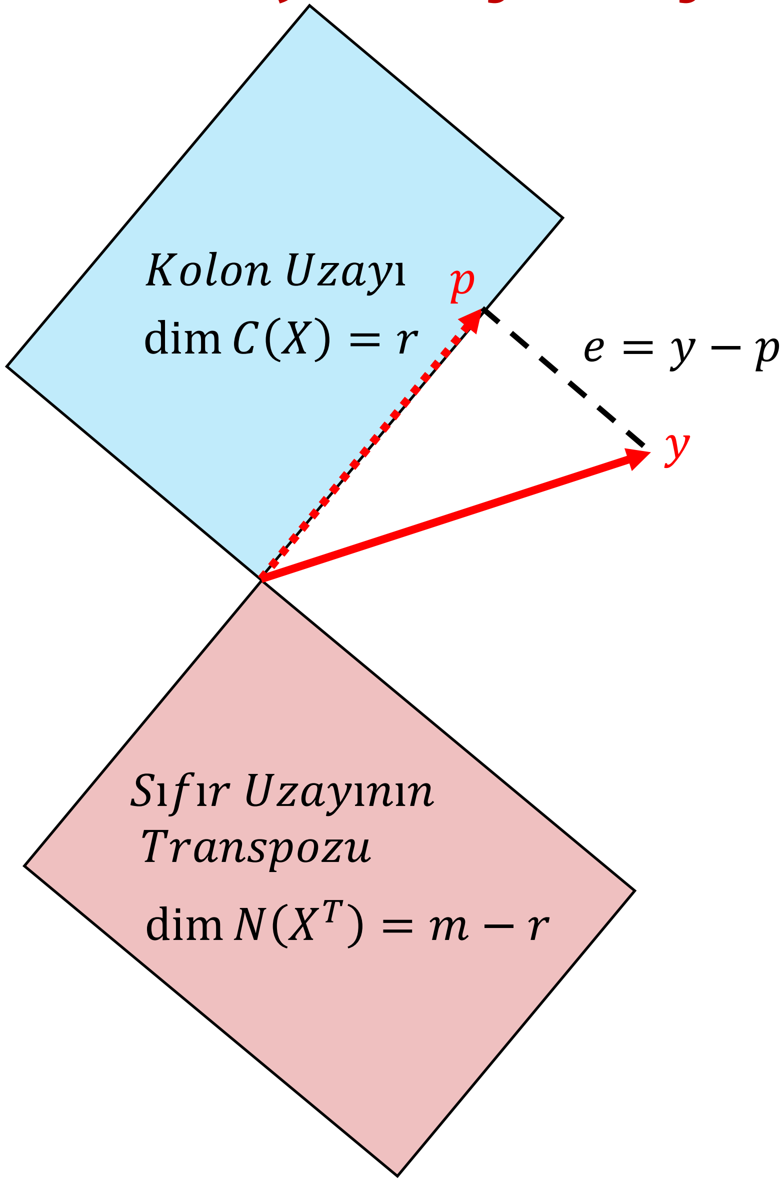
Neden Projeksiyon yaparız ?



- *Çünkü $Xw = y$ denkleminin çözümü olmayabilir.*
- *yani uygun y değerlerini alamayabilirim.*
- *Kolon uzayında çalışmam gerekli. Fakat kolon uzayında değilim.*
- *Sıfır uzayında da değilim. İkisinin arasında bir yerdeyim.*
- *Dolayısı ile kolon uzayında çalışmam gerektiğinden, kolon uzayına projeksiyon yapmalıyım.*
- *Bundan dolayı bizde; $X\hat{w} = p$ denklemini çözeriz.*
- *$\hat{w} \neq w$, çünkü $\hat{w}'$nin X matrisindeki kombinasyonları bize bir sonuç verir.*

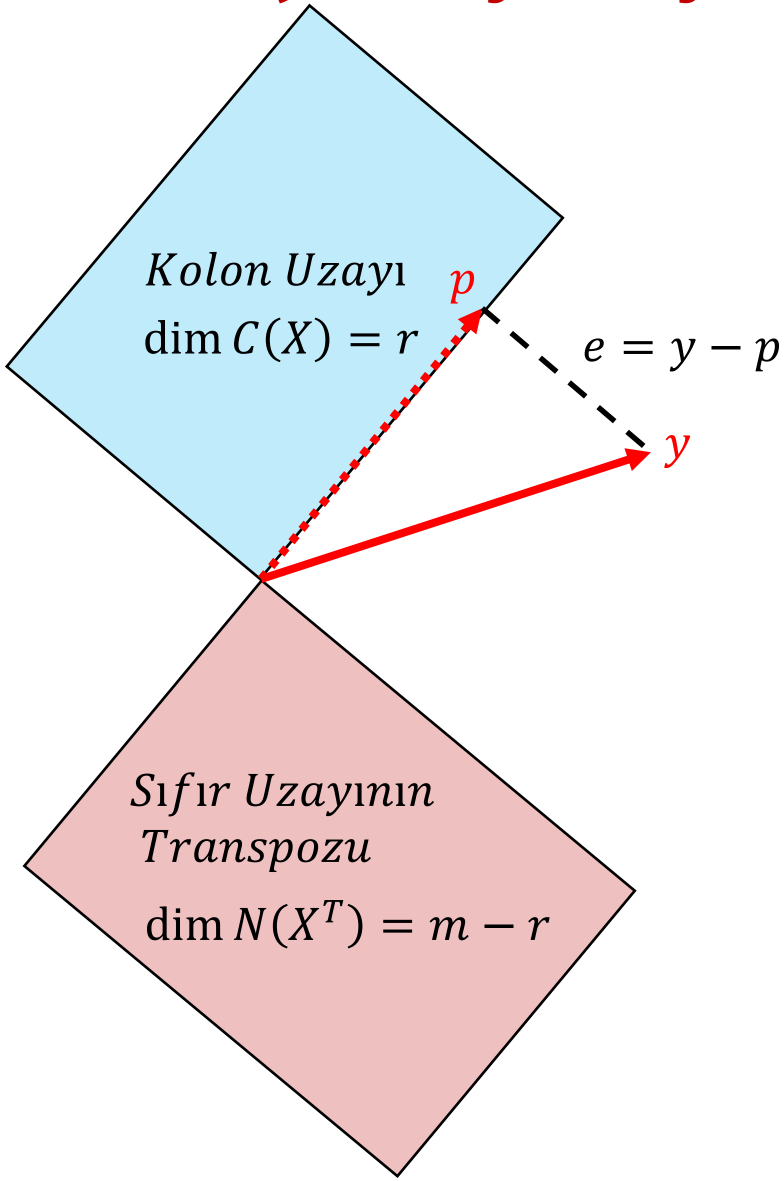
  (En iyi sonuç $\sim p = \hat{y}$)

Neden Projeksiyon yaparız ?



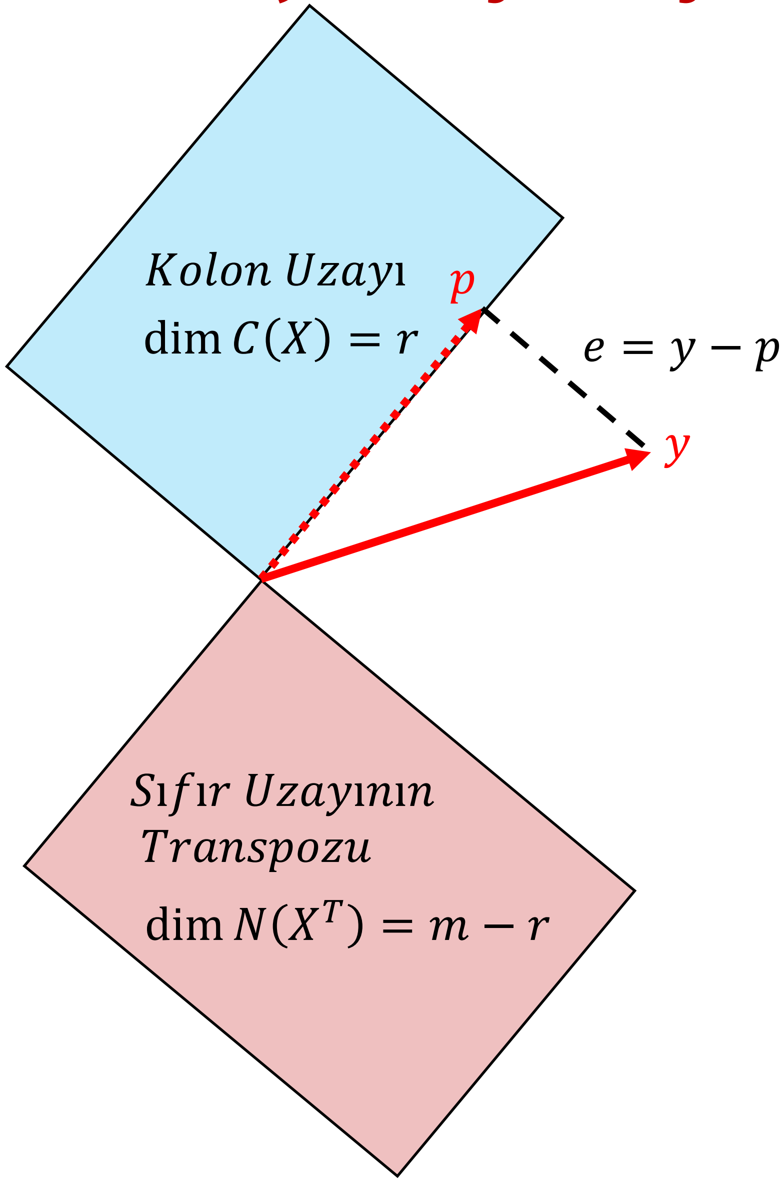
$$\boxed{P = \frac{xx^T}{x^Tx}}$$

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

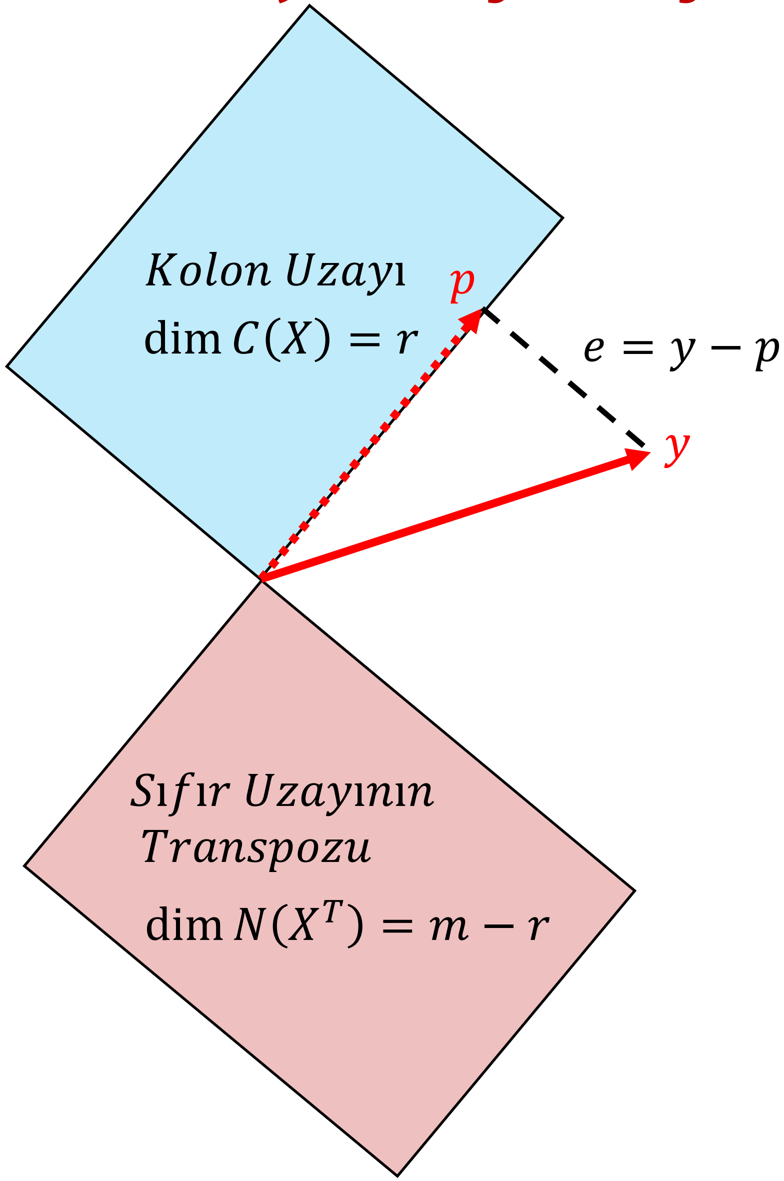
Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

$Çok\ boyuta\ geçildiğinde;$

Neden Projeksiyon yaparız ?

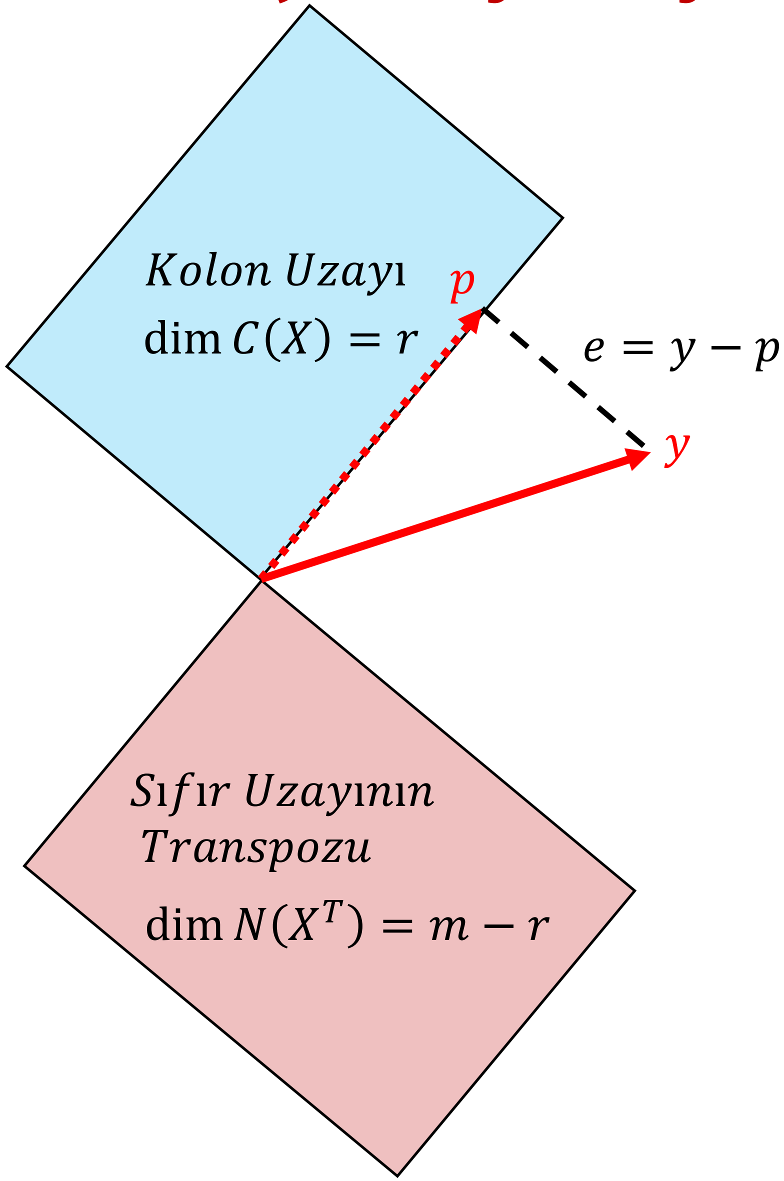


$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

Çok boyuta geçildiğinde;

$$e = y - p = y - \hat{y} = y - X\hat{w}$$

Neden Projeksiyon yaparız ?



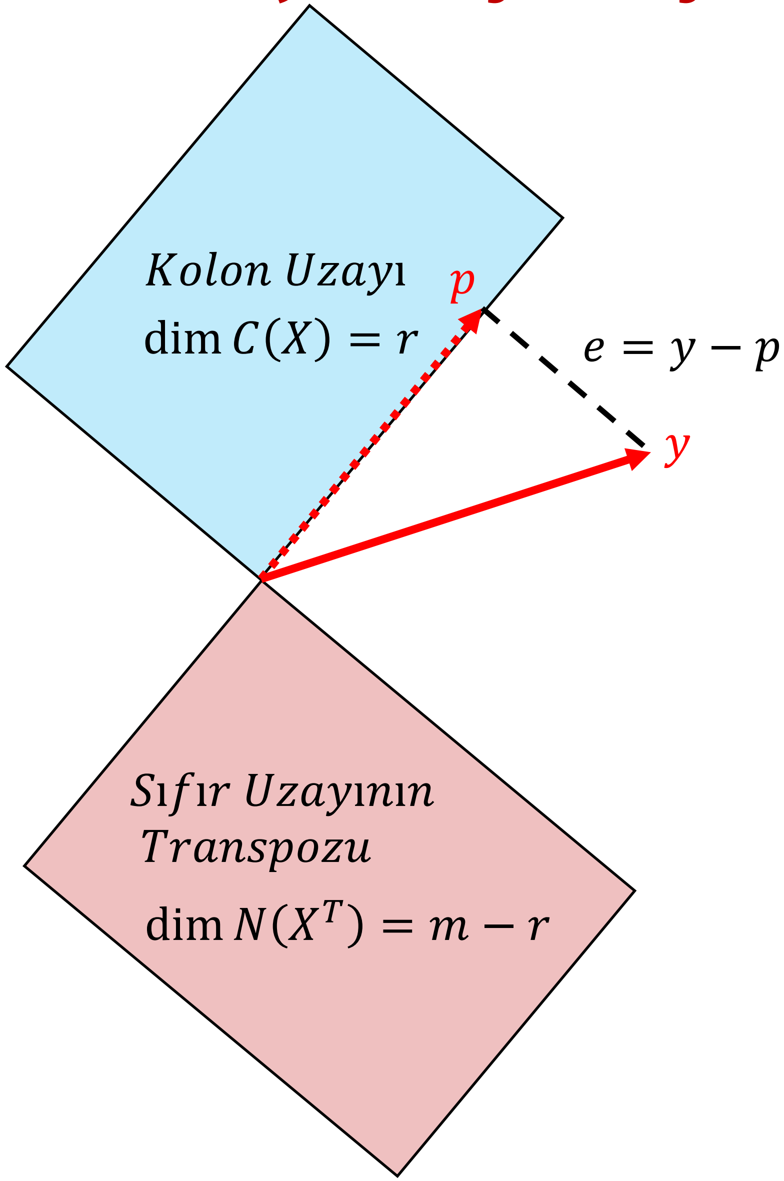
$$\boxed{P = \frac{xx^T}{x^T x}} \longrightarrow 1\ boyutta \sim 1d$$

Çok boyuta geçildiğinde;

$e = y - p = y - \hat{y} = y - X\hat{w}$

$X \perp e$

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

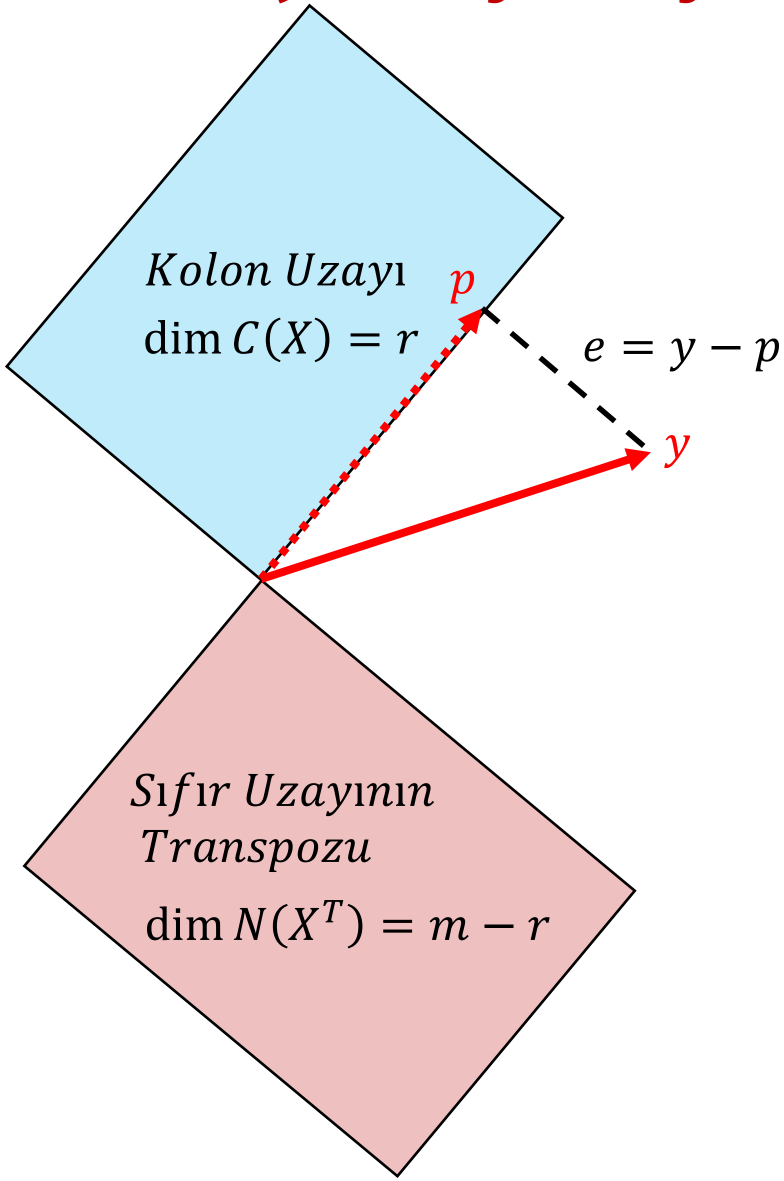
Çok boyuta geçildiğinde;

$e = y - p = y - \hat{y} = y - X\hat{w}$

$X \perp e$

$X^T(y - X\hat{w}) = 0$

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^T x}} \longrightarrow 1\ boyutta \sim 1d$$

Çok boyuta geçildiğinde;

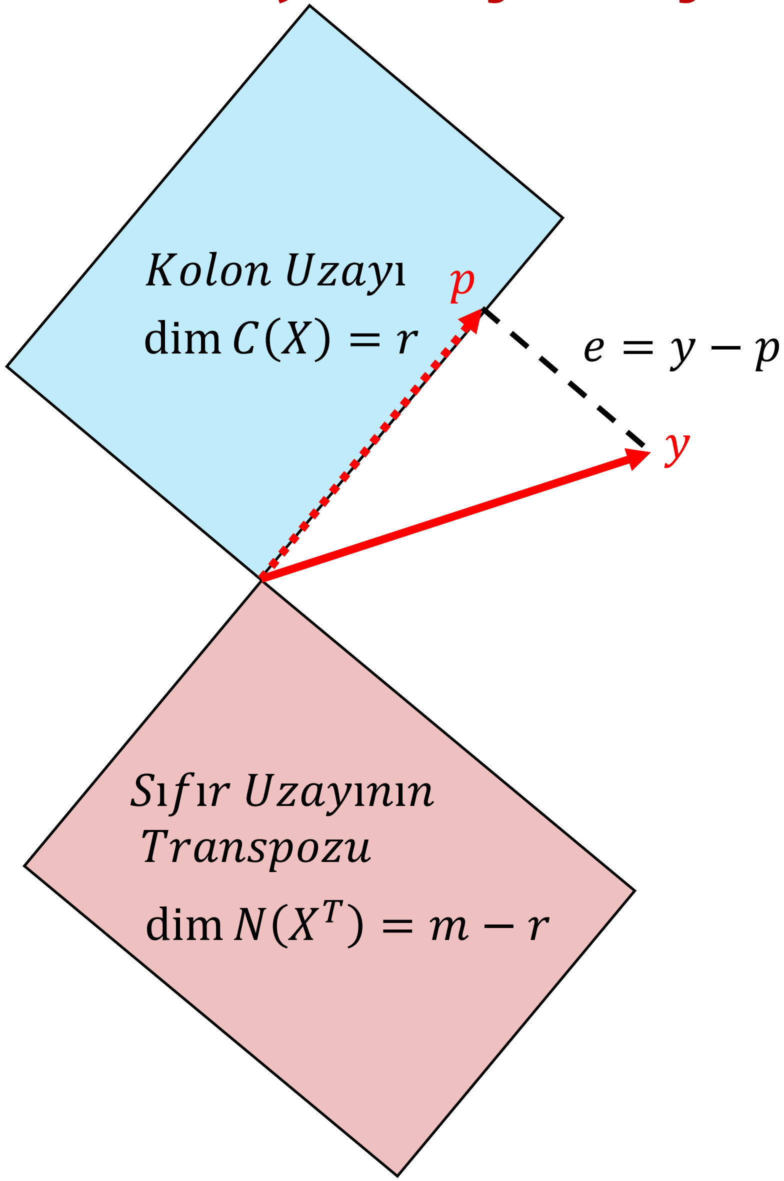
$$e = y - p = y - \hat{y} = y - X\widehat{w}$$

$$X \perp e$$

$$X^T(y - X\widehat{w}) = 0$$

$$X^T y - X^T X\widehat{w} = 0$$

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

$Çok\ boyuta\ geçildiğinde;$

$$e = y - p = y - \hat{y} = y - X\hat{w}$$

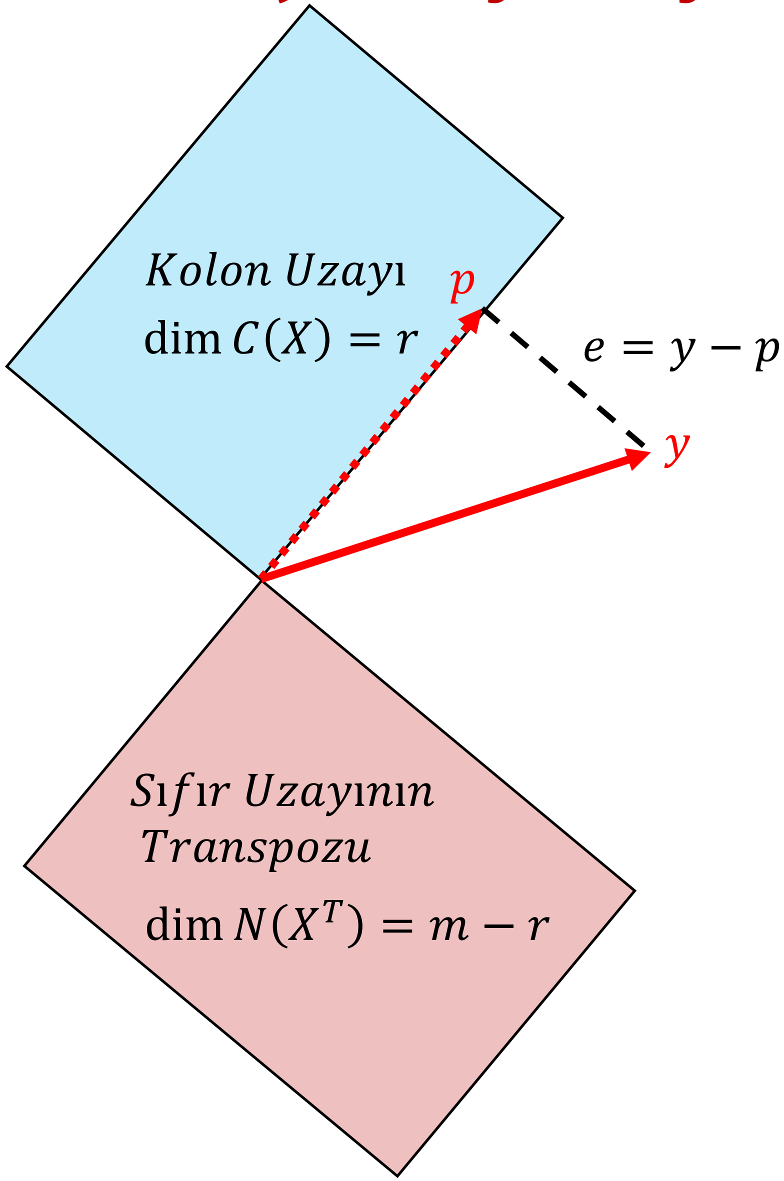
$$X \perp e$$

$$X^T(y - X\hat{w}) = 0$$

$$X^Ty - X^TX\hat{w} = 0$$

$$X^TX\hat{w} = X^Ty$$

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

$Çok\ boyuta\ geçildiğinde;$

$$e = y - p = y - \hat{y} = y - X\hat{w}$$

$$X \perp e$$

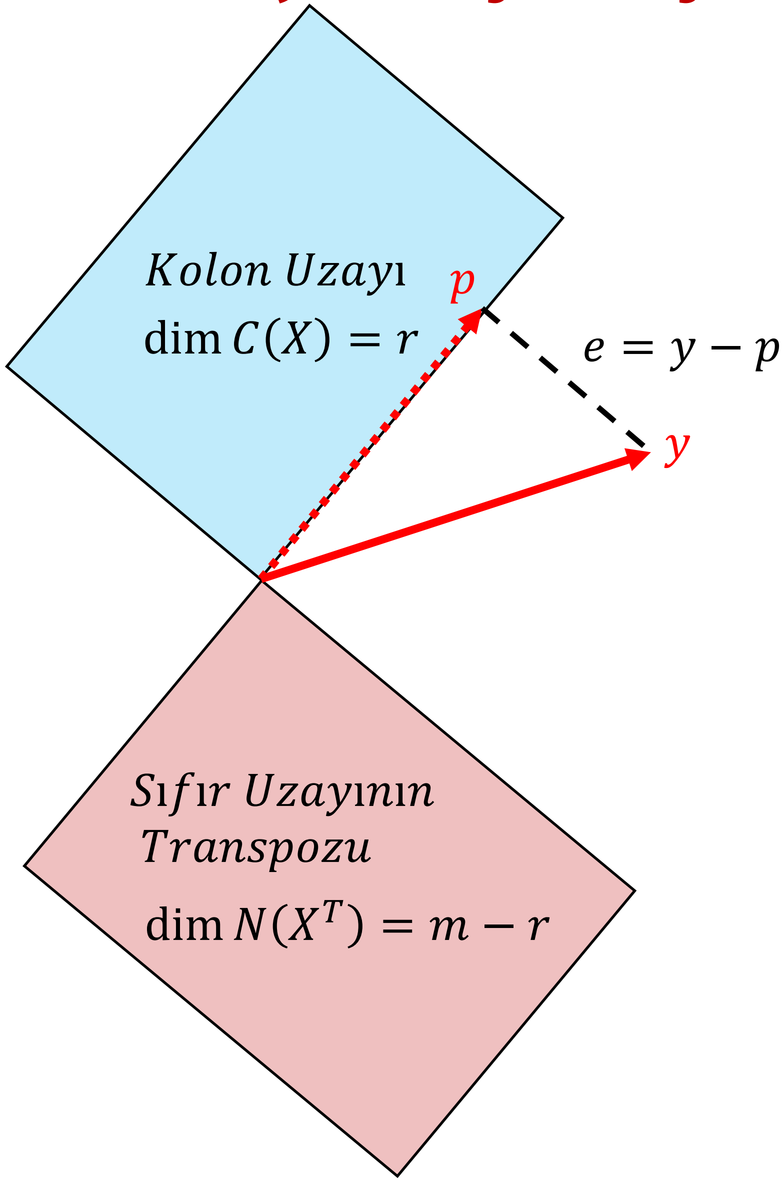
$$X^T(y - X\hat{w}) = 0$$

$$X^Ty - X^TX\hat{w} = 0$$

$$X^TX\hat{w} = X^Ty$$

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

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

Çok boyuta geçildiğinde;

$e = y - p = y - \hat{y} = y - X\hat{w}$

$X \perp e$

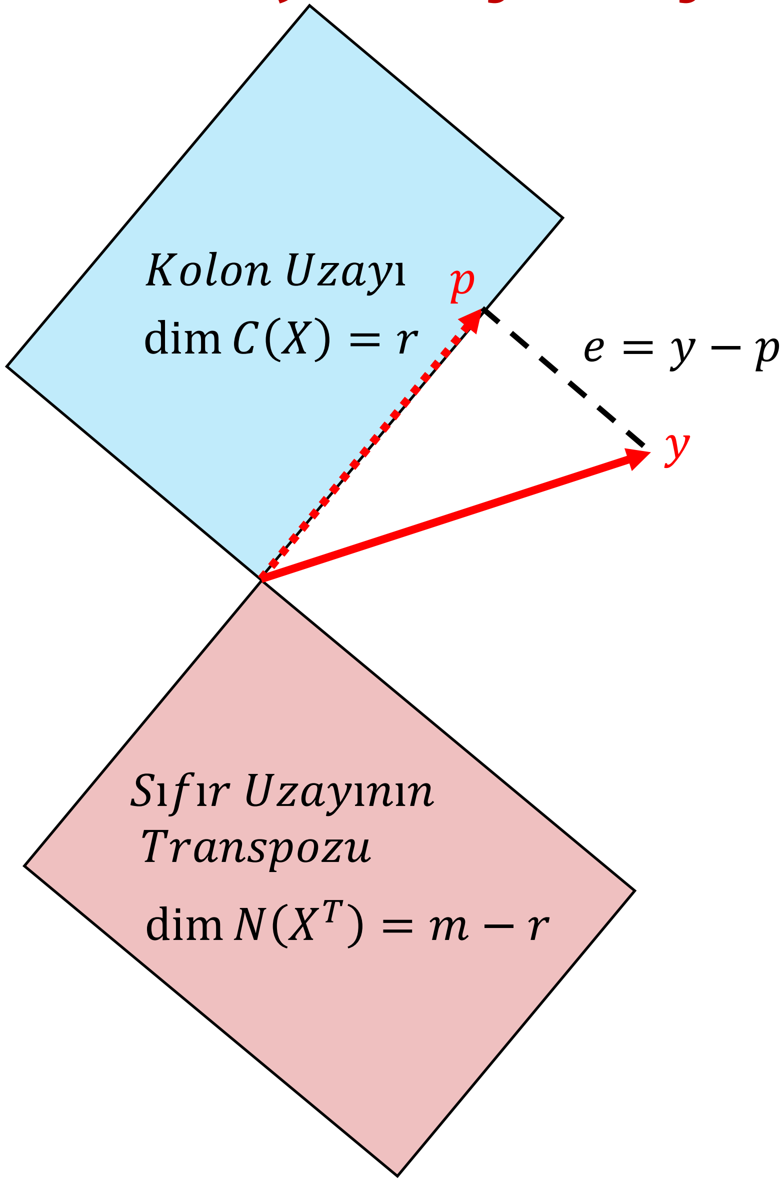
$X^T(y - X\hat{w}) = 0$

$X^Ty - X^TX\hat{w} = 0$

$X^TX\hat{w} = X^Ty$

$$\boxed{\hat{w} = (X^TX)^{-1}X^Ty}$$

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

$Çok\ boyuta\ geçildiğinde;$

$$e = y - p = y - \hat{y} = y - X\hat{w}$$

$$X \perp e$$

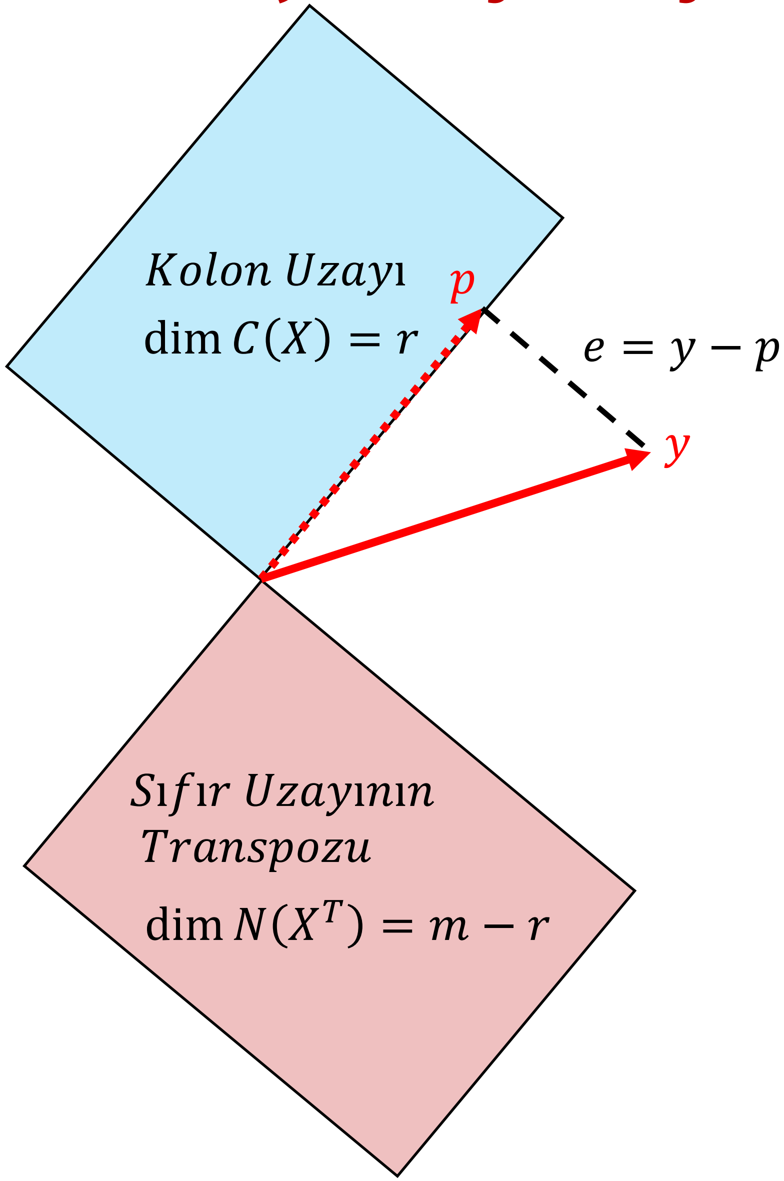
$$X^T(y - X\hat{w}) = 0$$

$$X^Ty - X^TX\hat{w} = 0$$

$$X^TX\hat{w} = X^Ty$$

$$\boxed{\hat{w} = (X^TX)^{-1}X^Ty} \longrightarrow Çok\ boyutta$$

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

$Çok\ boyuta\ geçildiğinde;$

$$e = y - p = y - \hat{y} = y - X\hat{w}$$

$$X \perp e$$

$$X^T(y - X\hat{w}) = 0$$

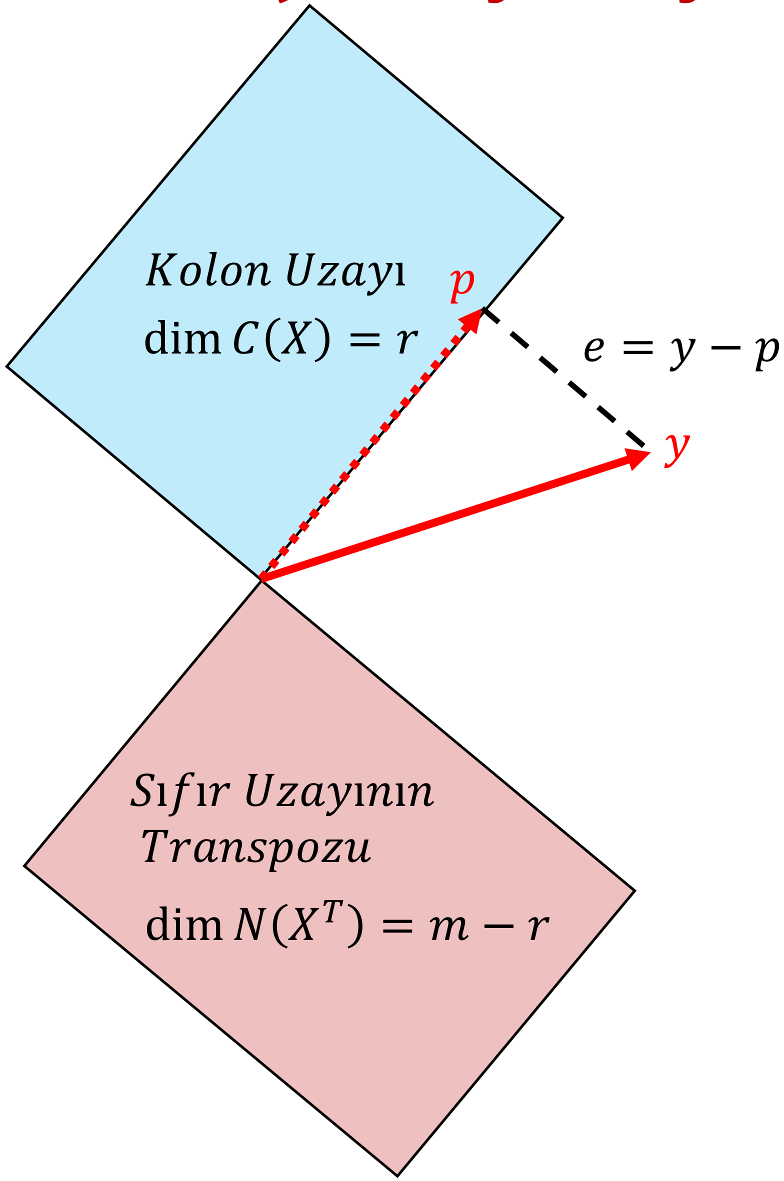
$$X^Ty - X^TX\hat{w} = 0$$

$$X^TX\hat{w} = X^Ty$$

$$\boxed{\hat{w} = (X^TX)^{-1}X^Ty} \longrightarrow Çok\ boyutta$$

$$p = X\hat{w} = X(X^TX)^{-1}X^Ty$$

Neden Projeksiyon yaparız ?



$$\boxed{P = \frac{xx^T}{x^Tx}} \longrightarrow 1\ boyutta \sim 1d$$

Çok boyuta geçildiğinde;

$$e = y - p = y - \hat{y} = y - X\hat{w}$$

$$X \perp e$$

$$X^T(y - X\hat{w}) = 0$$

$$X^Ty - X^TX\hat{w} = 0$$

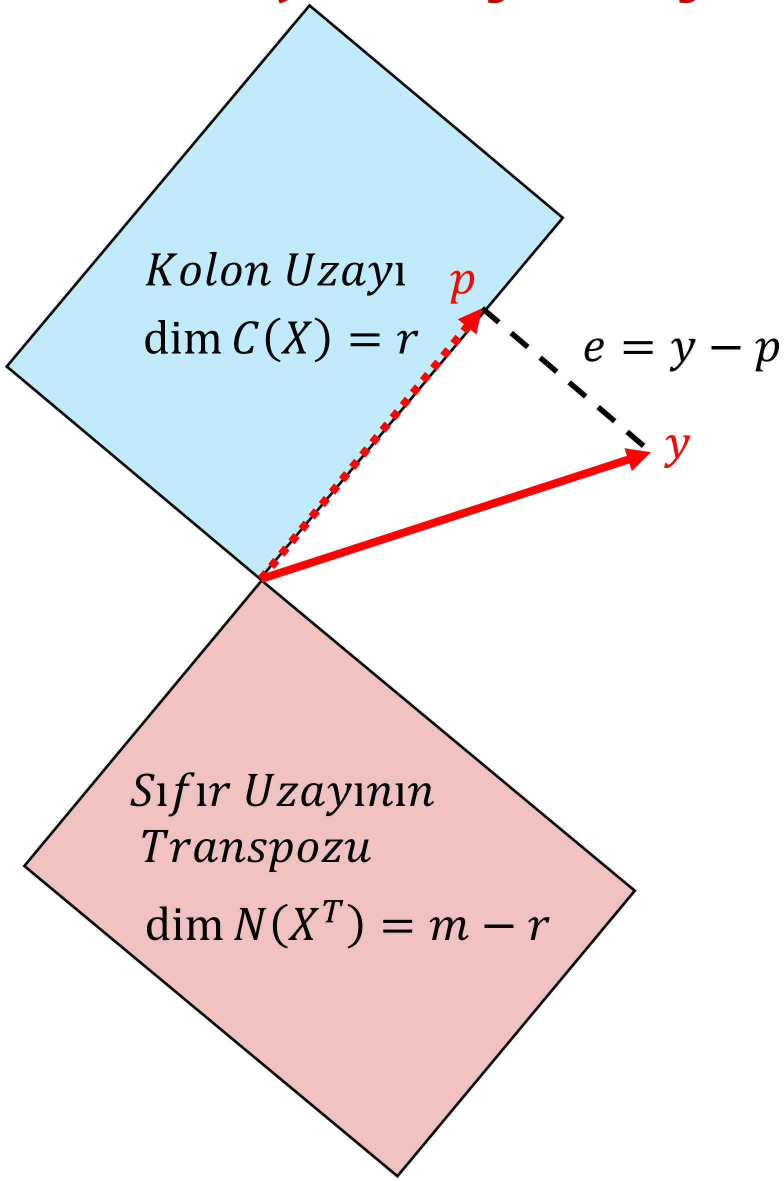
$$X^TX\hat{w} = X^Ty$$

$$\boxed{\hat{w} = (X^TX)^{-1}X^Ty} \longrightarrow Çok\ boyutta$$

$$p = X\hat{w} = X(X^TX)^{-1}X^Ty$$

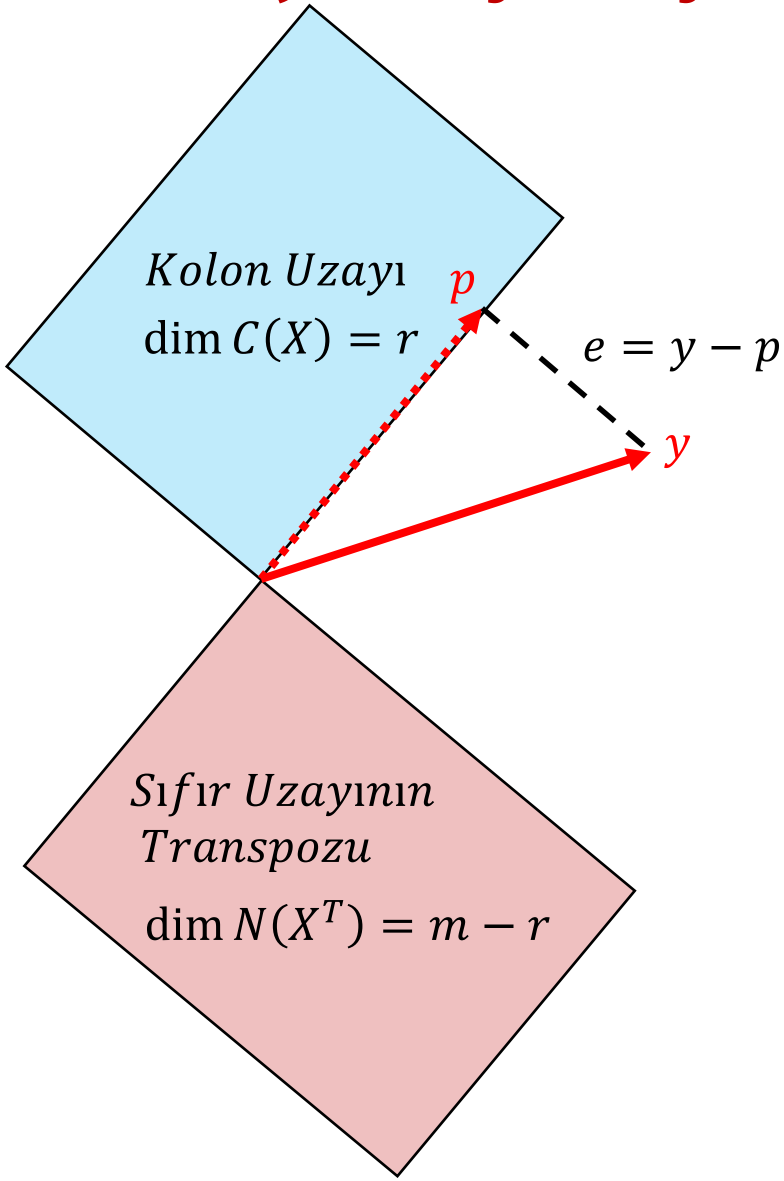
$$P = X(X^TX)^{-1}X^T$$

Neden Projeksiyon yaparız ?



- $P = X(X^TX)^{-1}X^T$

Neden Projeksiyon yaparız ?

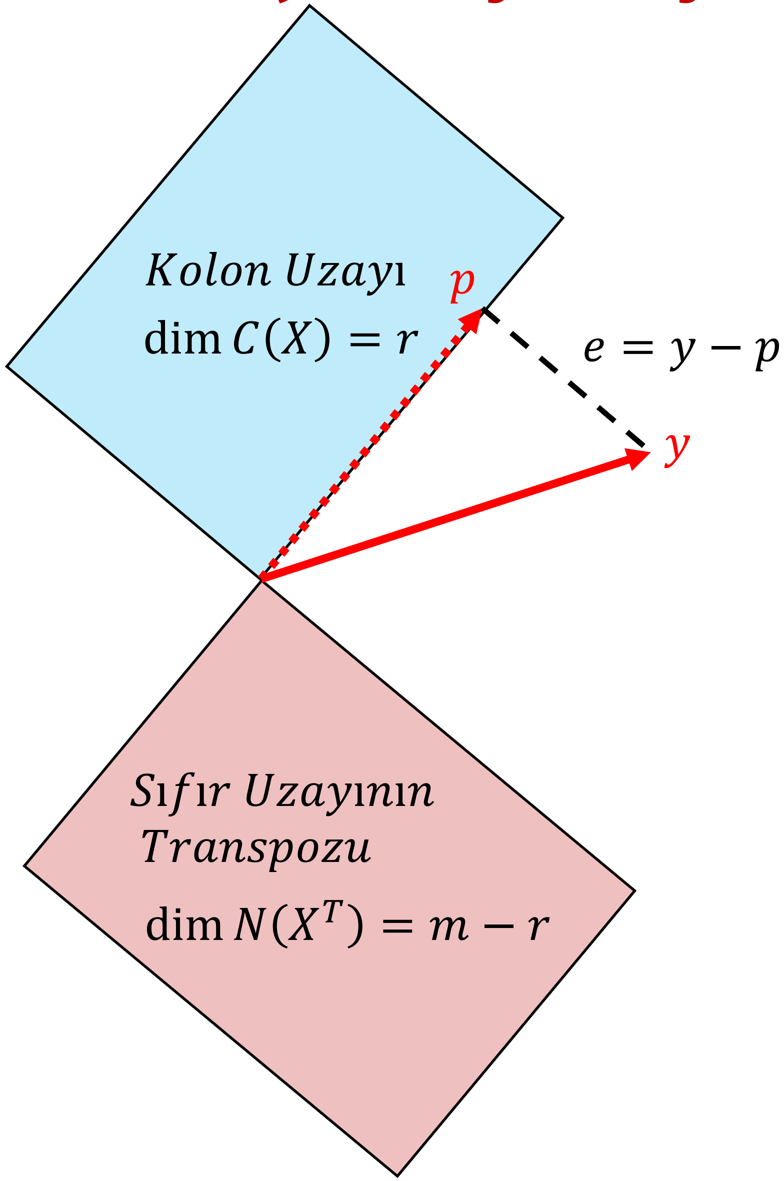


- $P = X(X^TX)^{-1}X^T$
- $Eğer\ tersi\ alınan\ değerin\ içini\ dışarıya\ dağıtır\ isek,$

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

$birim\ matrisini\ elde\ ederiz.$

Neden Projeksiyon yaparız ?



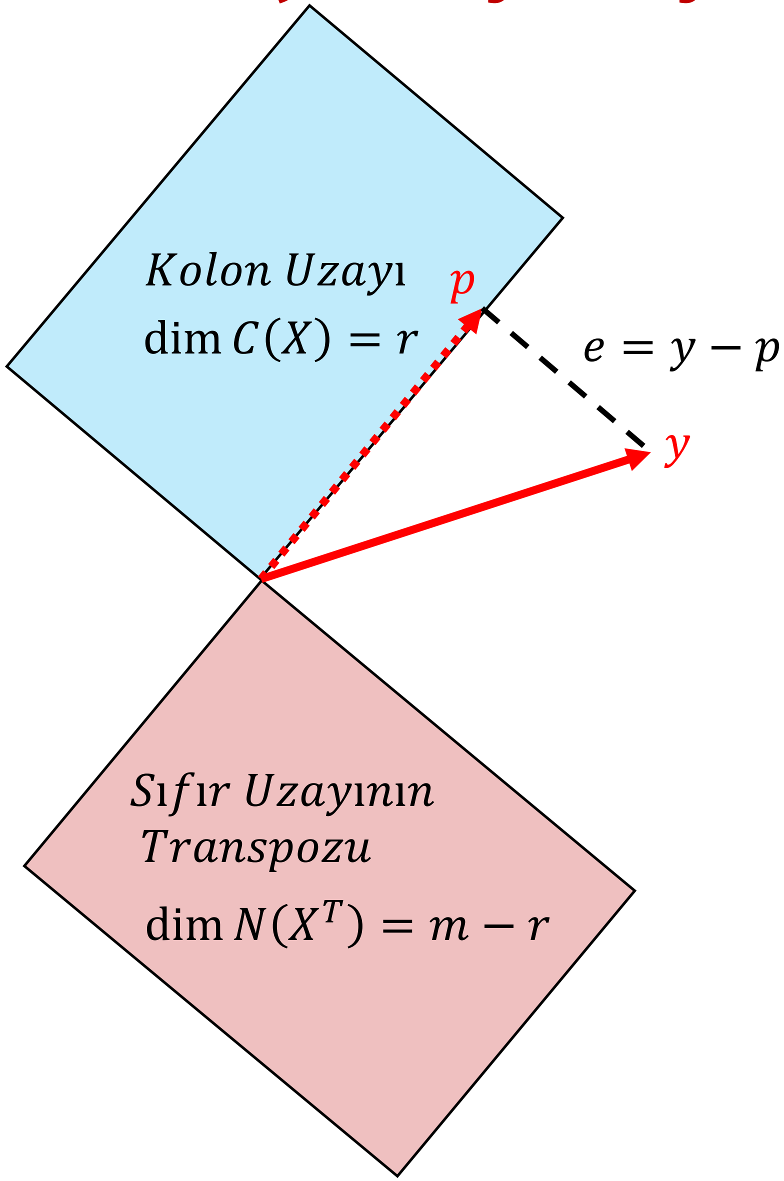
- $P = X(X^TX)^{-1}X^T$
- *Eğer tersi alınan değerin içini dışarıya dağıtır isek,*

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

birim matrisini elde ederiz.

- *Bu işlem kabul edilemez,*
  çünkü X matrisinin tersi alınamıyacak
  bir durumda olması gereklidir.

Neden Projeksiyon yaparız ?



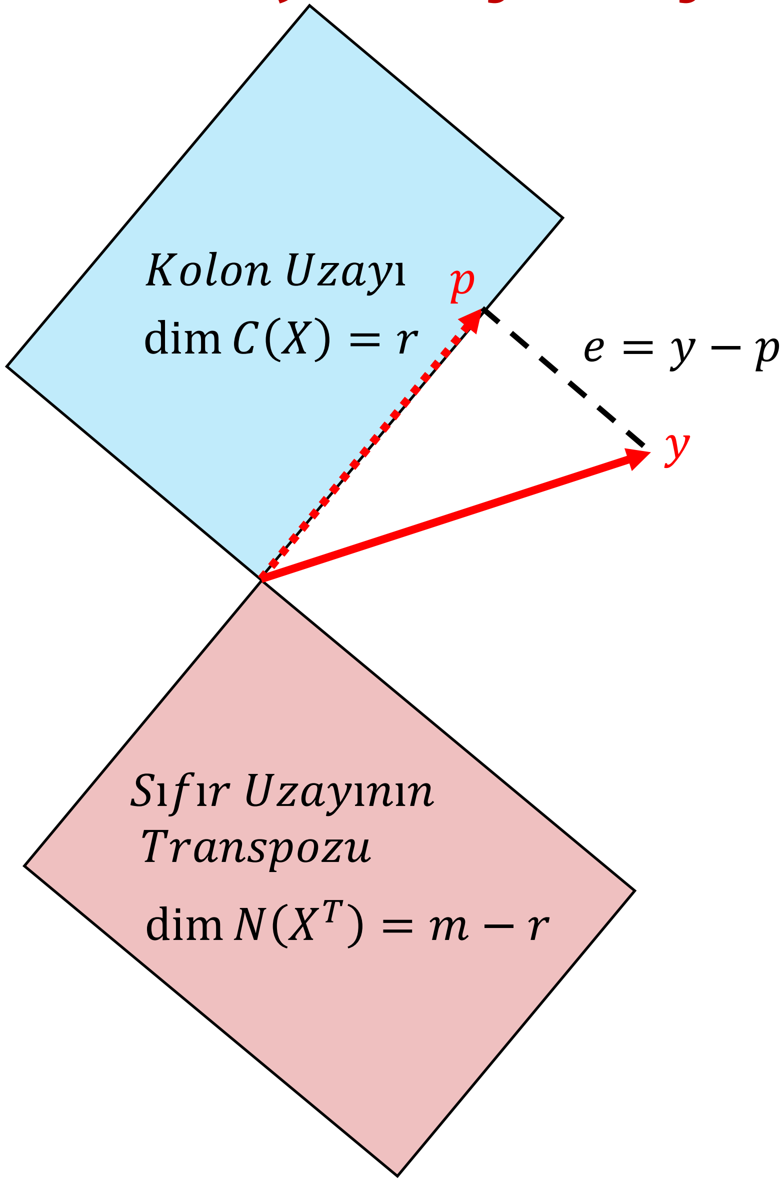
- $P = X(X^TX)^{-1}X^T$
- *Eğer tersi alınan değerin içini dışarıya dağıtır isek,*

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

  birim matrisini elde ederiz.

- *Bu işlem kabul edilemez, çünkü X matrisinin tersi alınamıyacak bir durumda olması gereklidir.*
- *Eğer X matrisi kare* ve *terslenebilir bir matris olsaydı, o zaman Projeksiyon matrisi Birim matris olurdu.*

Neden Projeksiyon yaparız ?



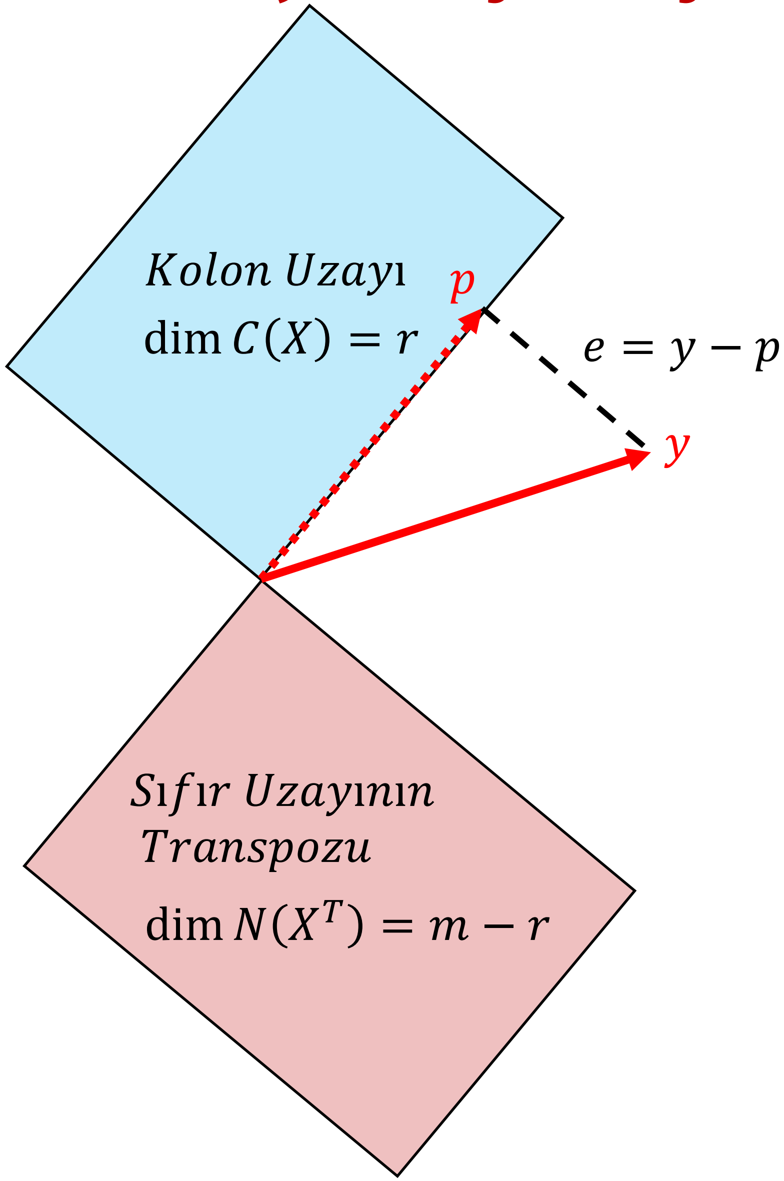
- $P = X(X^T X)^{-1} X^T$
- $Eğer\ tersi\ alınan\ değerin\ içini\ dışarıya\ dağıtır\ isek,$

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

$birim\ matrisini\ elde\ ederiz.$

- $Bu\ işlem\ kabul\ edilemez,$
  $çünkü\ X\ matrisinin\ tersi\ alınamıyacak$
  $bir\ durumda\ olması\ gereklidir.$
- $Eğer\ X\ matrisi\ kare$ ve $terslenebilir\ bir\ matris\ olsaydı,$
  $o\ zaman\ Projeksiyon\ matrisi\ Birim\ matris\ olurdu.$
- $Çok\ boyutta\ projeksiyon\ matrisinin$
  $transpozu\ ve\ tekrar\ çarpımı\ yine\ kendisine\ eşittir.$

Neden Projeksiyon yaparız ?



- $P = X(X^TX)^{-1}X^T$
- *Eğer tersi alınan değerin içini dışarıya dağıtır isek,*

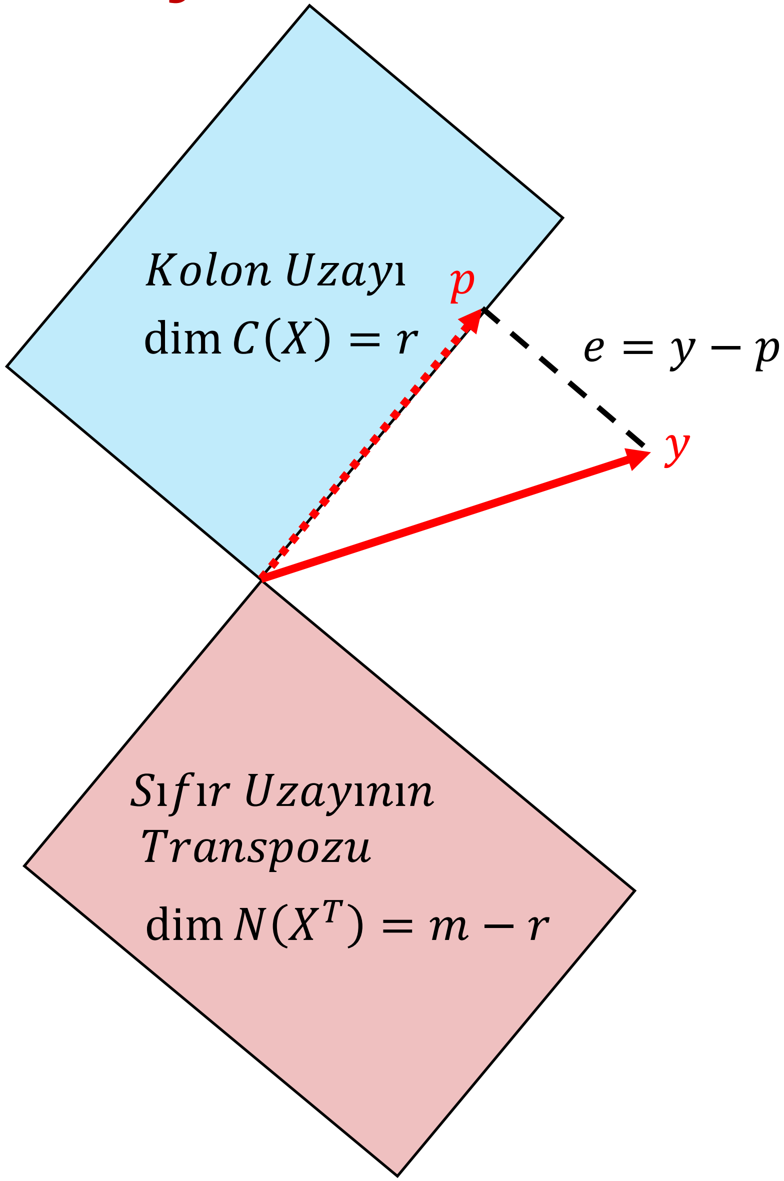
$$P = XX^{-1}(X^T)^{-1}X^T = I$$

birim matrisini elde ederiz.

- *Bu işlem kabul edilemez,*
  çünkü X matrisinin tersi alınamıyacak
  bir durumda olması gereklidir.
- *Eğer X matrisi kare* ve *terslenebilir bir matris olsaydı,*
  o zaman Projeksiyon matrisi Birim matris olurdu.
- *Çok boyutta projeksiyon matrisinin*
  transpozu ve tekrar çarpımı yine kendisine eşittir.

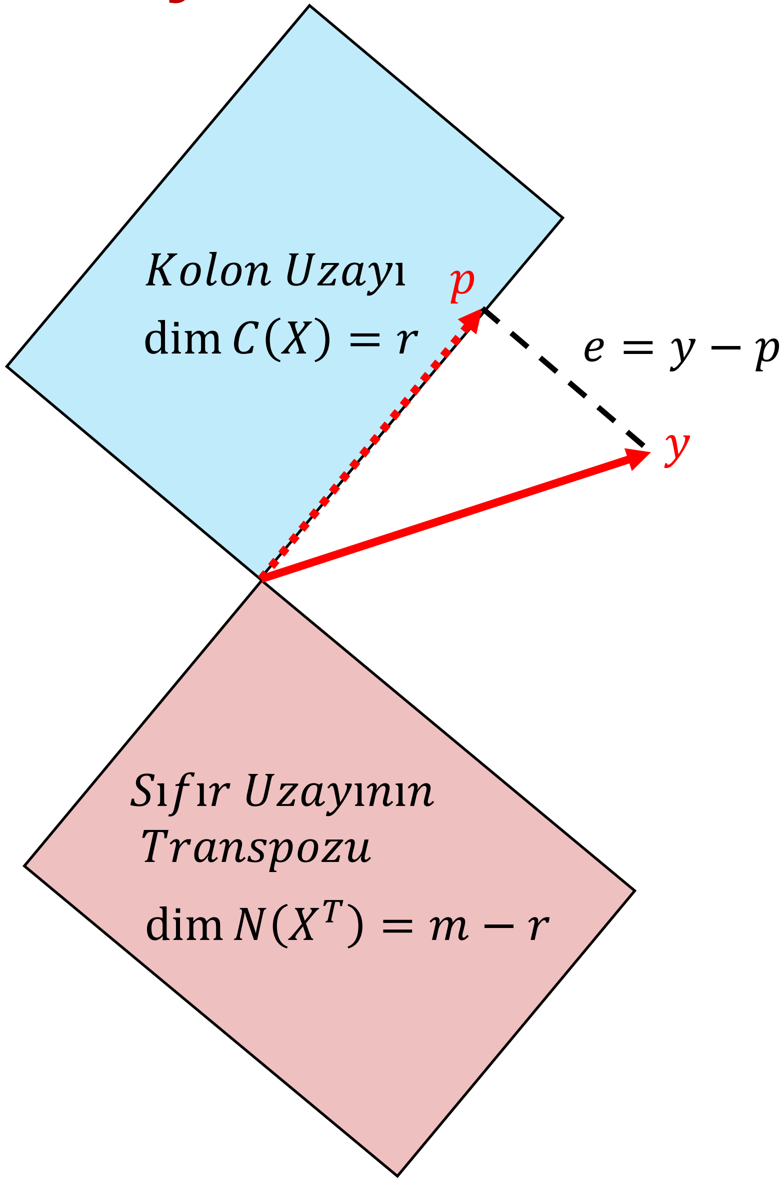
$$P^T = P, \qquad P^2 = P$$

Projeksiyon formulleri



$Aradığımız, elde\ edebileceğimiz\ en\ iyi\ sonuçtur$
$(makine\ öğrenmesi\ için\ ağırlık\ değerleri).$

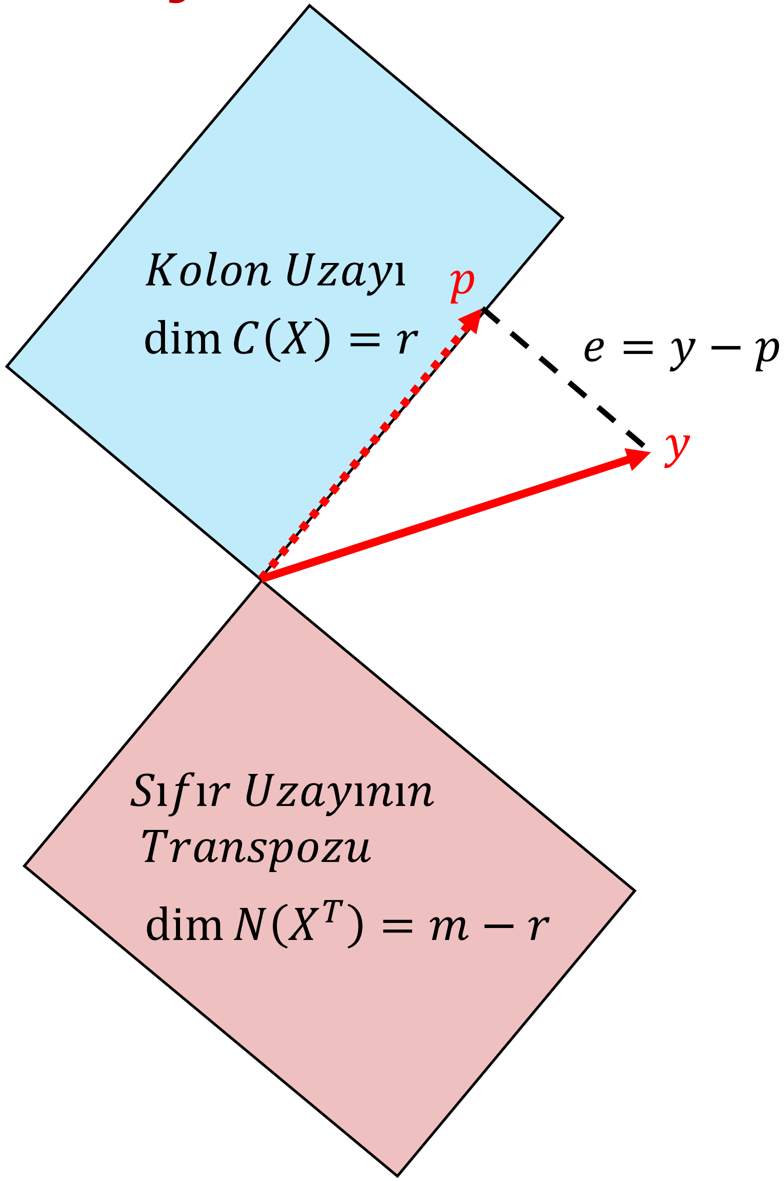
Projeksiyon formulleri



$Aradığımız, elde\ edebileceğimiz\ en\ iyi\ sonuçtur$
$(makine\ öğrenmesi\ için\ ağırlık\ değerleri).$

- $X^T X\hat{w} = X^T y,\ çözebileceğim\ denklem$

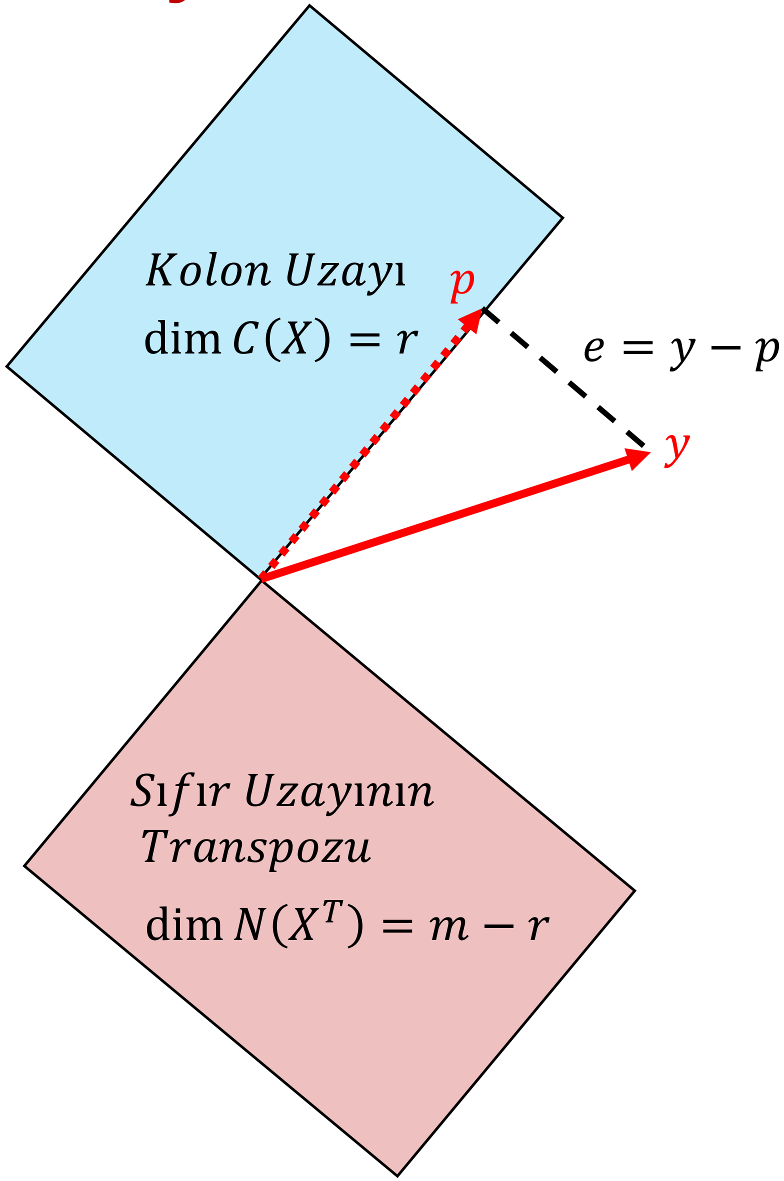
Projeksiyon formulleri



$Aradığımız, elde\ edebileceğimiz\ en\ iyi\ sonuçtur$ $(makine\ öğrenmesi\ için\ ağırlık\ değerleri).$

- $X^T X\hat{w} = X^T y,$ $çözebileceğim\ denklem$
- $\hat{w} = (X^T X)^{-1} X^T y,$ $en\ iyi\ w\ (en\ yakın\ w)$

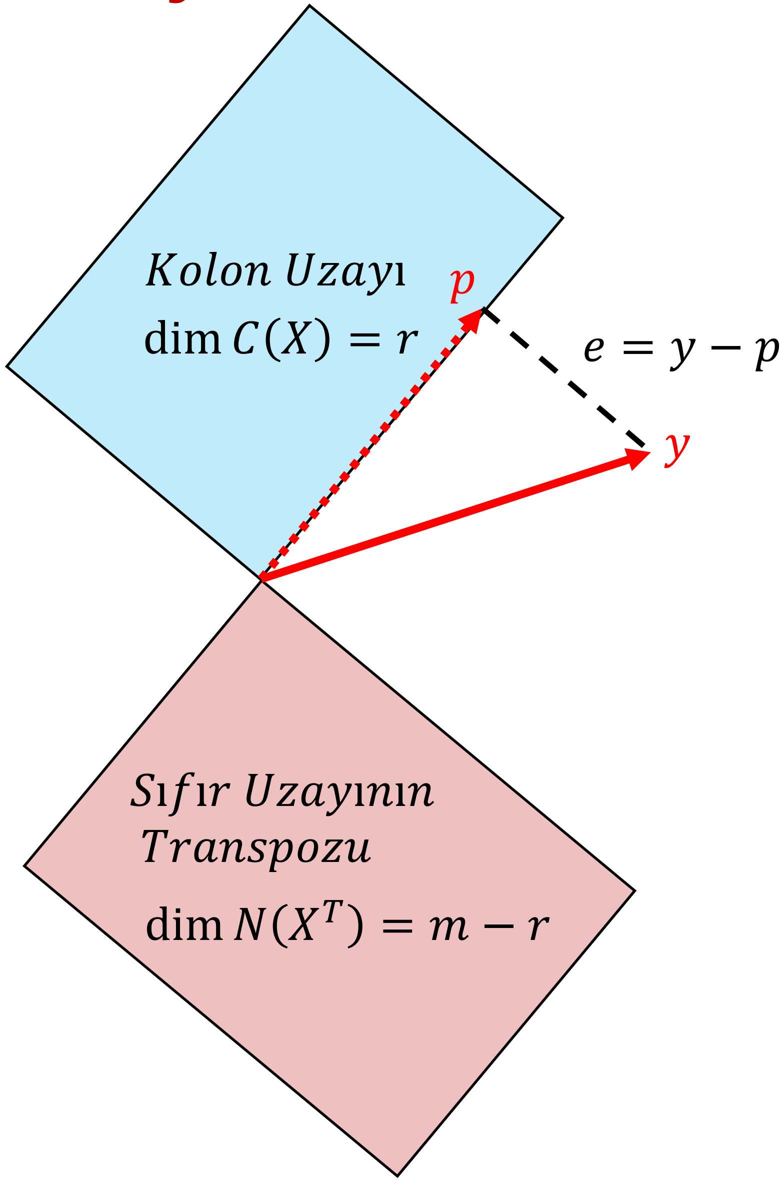
Projeksiyon formulleri



Aradığımız, elde edebileceğimiz en iyi sonuçtur (makine öğrenmesi için ağırlık değerleri).

- $X^T X\hat{w} = X^T y$, *çözebileceğim denklem*
- $\hat{w} = (X^T X)^{-1} X^T y$, *en iyi w (en yakın w)*
- $p = X(X^T X)^{-1} X^T y$, *en iyi projeksiyon*

Projeksiyon formulleri



$Aradığımız, elde\ edebileceğimiz\ en\ iyi\ sonuçtur$
$(makine$ öğrenmesi $için\ ağırlık\ değerleri).$

- $X^T X\widehat{w} = X^T y,$ çözebileceğim $denklem$
- $\widehat{w} = (X^T X)^{-1} X^T y,\ en\ iyi\ w\ (en\ yakın\ w)$
- $p = X(X^T X)^{-1} X^T y,\ en\ iyi\ projeksiyon$
- $P = X(X^T X)^{-1} X^T,\ en\ iyi\ projeksiyon\ matrisi$

Pseudo inverse

Denklemimizde tersi alınamayan bir matris var ise $(X^T X)^{-1}$

"pseudoinverse" kullanılır.

https://en.wikipedia.org/wiki/Moore%E2%80%93Penrose_inverse

ilgili linkin QR kodu

$$\hat{w} = pseudoinverse(X^T X)X^T y$$

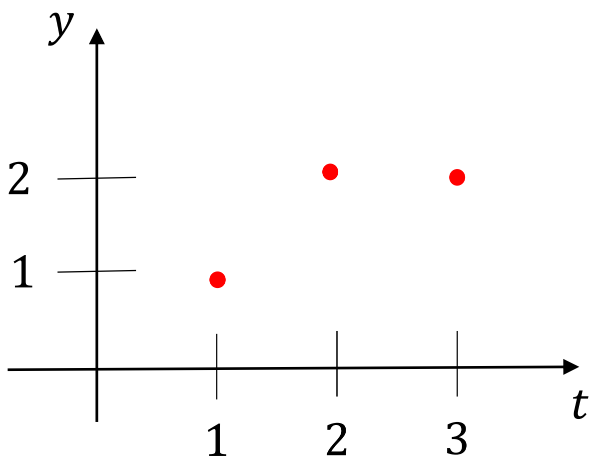
https://www.youtube.com/watch?v=Go2aLo7ZOlU

ilgili linkin QR kodu

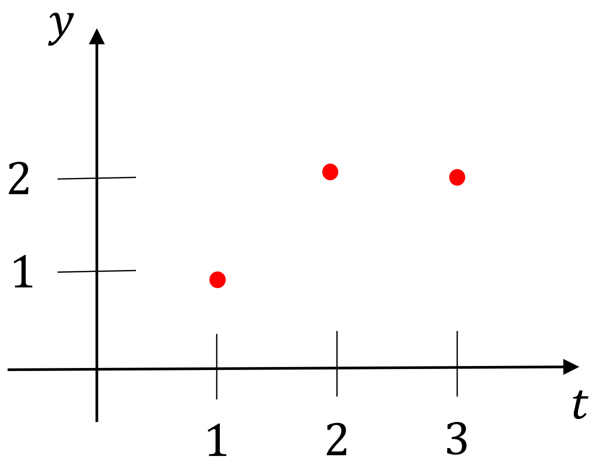
https://docs.scipy.org/doc/numpy/reference/generated/numpy.linalg.pinv.html#:~:text=pinv,-numpy.linalg.&text=Compute%20the%20(Moore%2DPenrose),including%20all%20large%20singular%20values.&text=Matrix%20or%20stack%20of%20matrices%20to%20be%20pseudo%2Dinverted.

ilgili linkin QR kodu

Örnek



Örnek

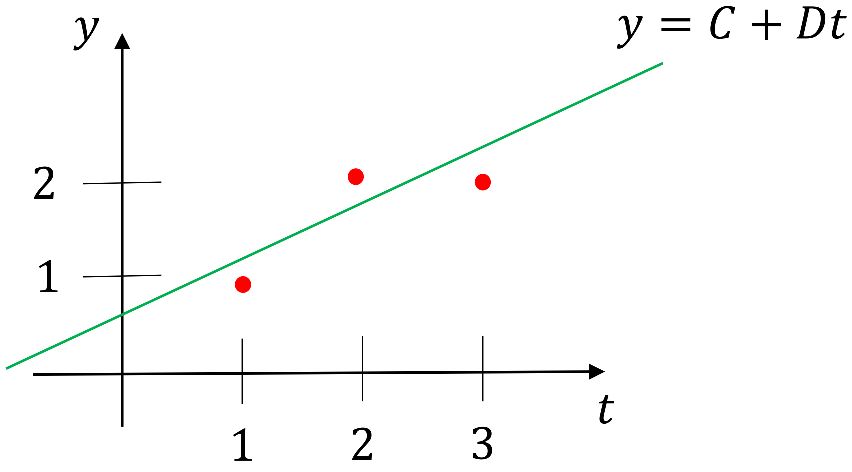


$$C + D = 1$$

$$C + 2D = 2$$

$$C + 3D = 2$$

Örnek

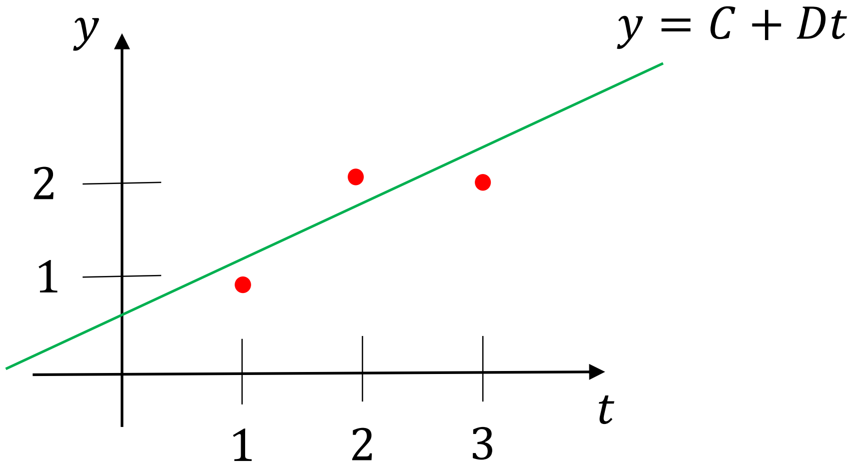


$$C + D = 1$$

$$C + 2D = 2$$

$$C + 3D = 2$$

Örnek



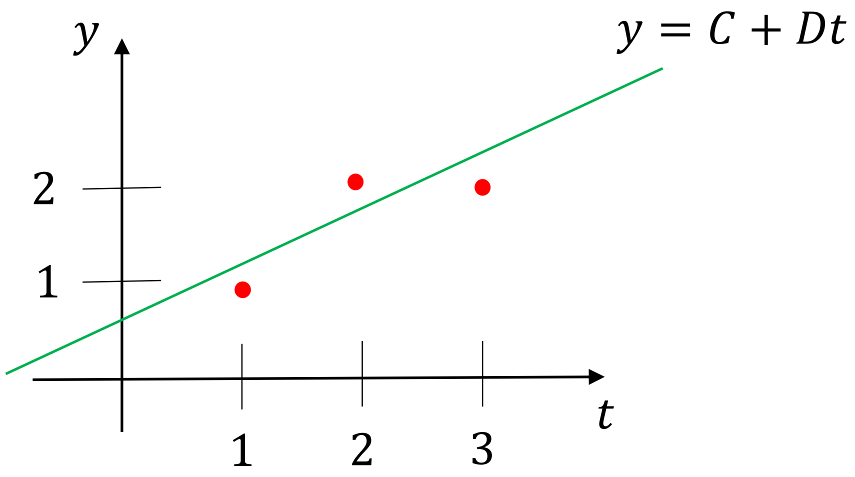
$$C + D = 1$$

$$C + 2D = 2$$

$$C + 3D = 2$$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{y}$$

Örnek



$$C + D = 1$$

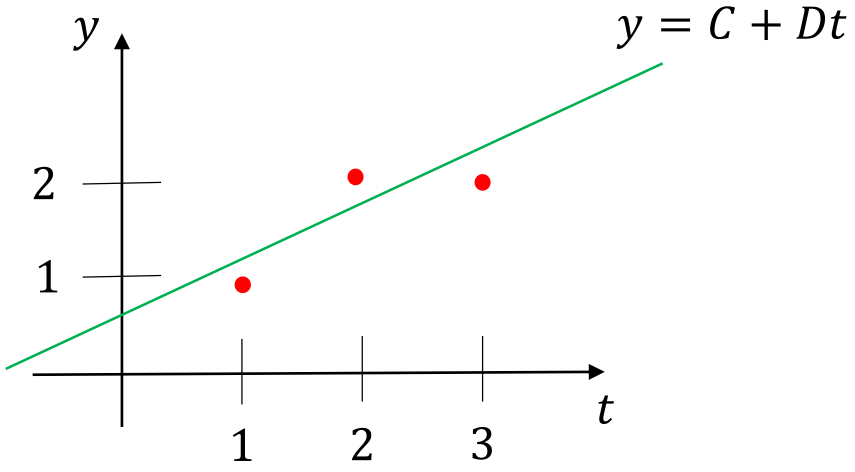
$$C + 2D = 2$$

$$C + 3D = 2$$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{y}$$

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

Örnek



$$C + D = 1$$

$$C + 2D = 2$$

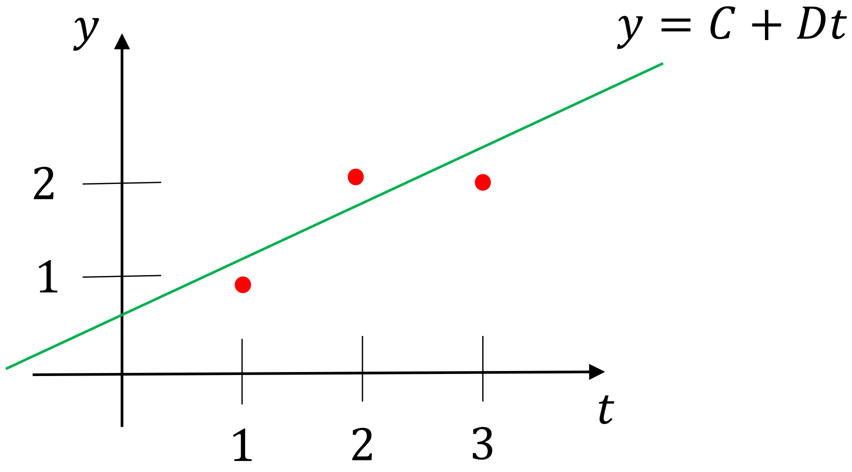
$$C + 3D = 2$$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{y}$$

$$\widehat{w} = (X^T X)^{-1} X^T y$$

$$\widehat{w} = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix}$$

Örnek



$$C + D = 1$$

$$C + 2D = 2$$

$$C + 3D = 2$$

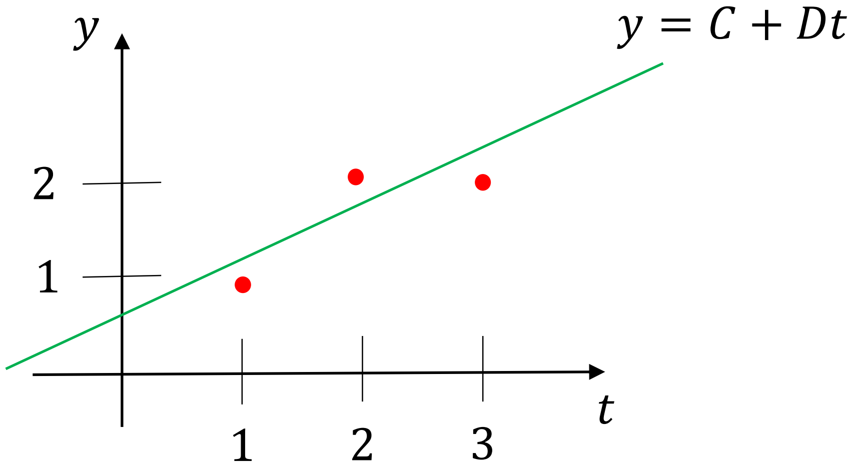
$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{y}$$

$$\widehat{w} = (X^T X)^{-1} X^T y$$

$$\widehat{w} = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix}$$

$$\hat{y} = C + Dt = \frac{2}{3} + \frac{1}{2} t$$

Örnek



$C + D = 1$

$C + 2D = 2$

$C + 3D = 2$

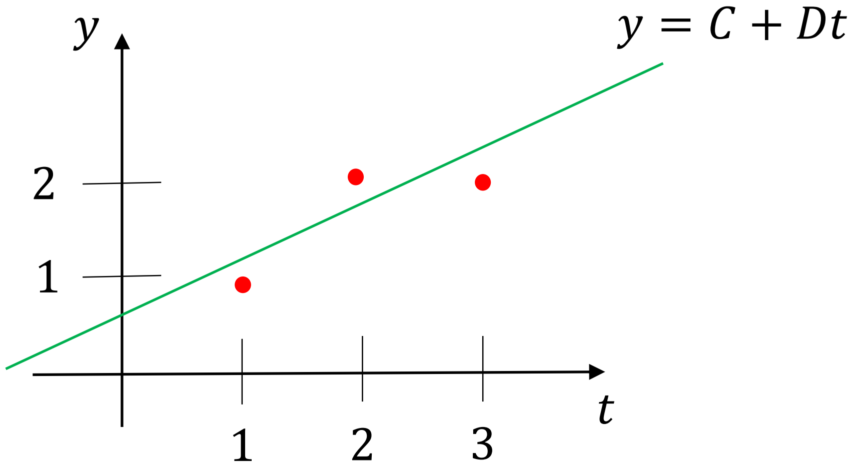
$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{y}$$

$$\hat{w} = (X^T X)^{-1} X^T y \qquad \hat{w} = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix}$$

$$\hat{y} = C + Dt = \frac{2}{3} + \frac{1}{2} t$$

$$\hat{y} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} + 1\frac{1}{2} \\ \frac{2}{3} + 2\frac{1}{2} \\ \frac{2}{3} + 3\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{7}{6} \\ \frac{5}{3} \\ \frac{13}{6} \end{bmatrix}$$

Örnek



$$C + D = 1$$
$$C + 2D = 2$$
$$C + 3D = 2$$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{y}$$

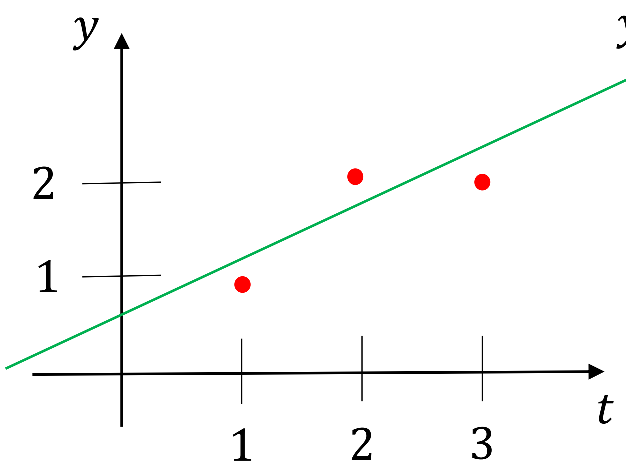
$$\widehat{w} = (X^T X)^{-1} X^T y \qquad \widehat{w} = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix}$$

$$\hat{y} = C + Dt = \frac{2}{3} + \frac{1}{2} t$$

$$e = y - \hat{y} = y - p$$

$$\hat{y} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} + 1\,\frac{1}{2} \\ \frac{2}{3} + 2\,\frac{1}{2} \\ \frac{2}{3} + 3\,\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{7}{6} \\ \frac{5}{3} \\ \frac{13}{6} \end{bmatrix}$$

Örnek



$$y = C + Dt$$

$$C + D = 1$$

$$C + 2D = 2$$

$$C + 3D = 2$$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{y}$$

$$\hat{w} = (X^T X)^{-1} X^T y \qquad \hat{w} = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix}$$

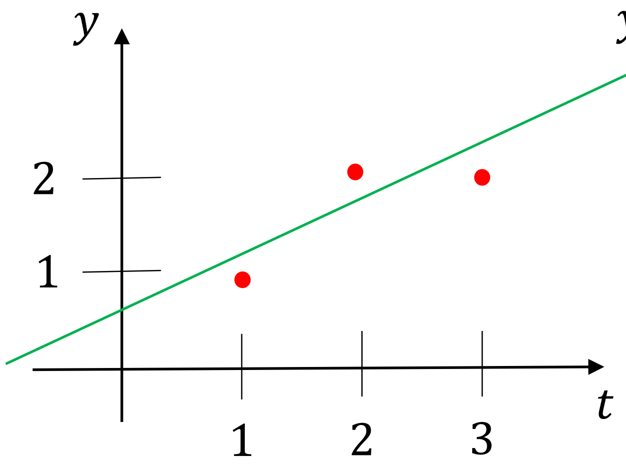
$$\hat{y} = C + Dt = \frac{2}{3} + \frac{1}{2} t$$

$$\hat{y} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} + 1\frac{1}{2} \\ \frac{2}{3} + 2\frac{1}{2} \\ \frac{2}{3} + 3\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{7}{6} \\ \frac{5}{3} \\ \frac{13}{6} \end{bmatrix}$$

$$e = y - \hat{y} = y - p$$

$$e = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{7}{6} \\ \frac{5}{3} \\ \frac{13}{6} \end{bmatrix} = \begin{bmatrix} \frac{-1}{6} \\ \frac{1}{3} \\ \frac{-1}{6} \end{bmatrix}$$

Örnek



$$y = C + Dt$$

$$C + D = 1$$

$$C + 2D = 2$$

$$C + 3D = 2$$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{y}$$

$$\hat{w} = (X^T X)^{-1} X^T y \qquad \hat{w} = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix}$$

$$\hat{y} = C + Dt = \frac{2}{3} + \frac{1}{2} t$$

$$\hat{y} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} + 1\frac{1}{2} \\ \frac{2}{3} + 2\frac{1}{2} \\ \frac{2}{3} + 3\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{7}{6} \\ \frac{5}{3} \\ \frac{13}{6} \end{bmatrix}$$

$$e = y - \hat{y} = y - p$$

$$e = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{7}{6} \\ \frac{5}{3} \\ \frac{13}{6} \end{bmatrix} = \begin{bmatrix} \frac{-1}{6} \\ \frac{1}{3} \\ \frac{-1}{6} \end{bmatrix}$$

$$\boxed{RMSE : \frac{\|e\|}{m} = \frac{1}{3}\sqrt{\frac{1}{36} + \frac{1}{9} + \frac{1}{36}} = \frac{1}{3}\sqrt{\frac{1}{6}}}$$