

# Makine Öğrenmesi

## Doğrusal Bağlanım

**İlker Birbil ve Utku Karaca**

Erasmus Üniversitesi Rotterdam

İstanbul'da Makine Öğrenmesi

27 Ocak – 2 Şubat, 2020



# Makine Öğrenmesi

```
graph TD; A[Makine Öğrenmesi] --> B[Doğrusal Bağlanım]; A --> C[Boyut Küçültme ve Düzenleştirme]; A --> D[Tekrar Örnekleme ve Model Değerlendirme]; A --> E[Sınıflandırma ve Ağaçlar]; A --> F[Güdümsüz Öğrenme]; A --> G[Yapay Sinir Ağları ve Derin Öğrenme];
```

Doğrusal Bağlanım

Boyut Küçültme  
ve  
Düzenleştirme

Tekrar Örnekleme  
ve  
Model Değerlendirme

Sınıflandırma  
ve  
Ağaçlar

Güdümsüz  
Öğrenme

Yapay Sinir Ağları  
ve  
Derin Öğrenme

$$Y = f(X) + \epsilon \xrightarrow{\text{yaklaşık?}} \hat{Y} = \hat{f}(X)$$

$$Y \approx \hat{f}(X)$$

### Basit Bağlanım (Simple Regression)

$$Y \approx \beta_0 + \beta_1 X$$

$$\text{satış} \approx \beta_0 + \beta_1 \text{TV}$$

### Çoklu Bağlanım (Multiple Regression)

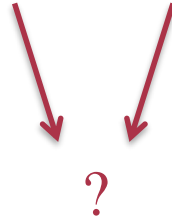
$$Y \approx \beta_0 + \beta_1 X_1 + \beta_1 X_2 + \cdots + \beta_p X_p$$

$$\text{satış} \approx \beta_0 + \beta_1 \text{TV} + \beta_2 \text{radyo} + \cdots + \beta_p \text{gazete}$$

# Basit Bağlanım

$$Y \approx \beta_0 + \beta_1 X$$

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$



eğitim verisi

$$\{(x_i, y_i) : 1, \dots, n\}$$

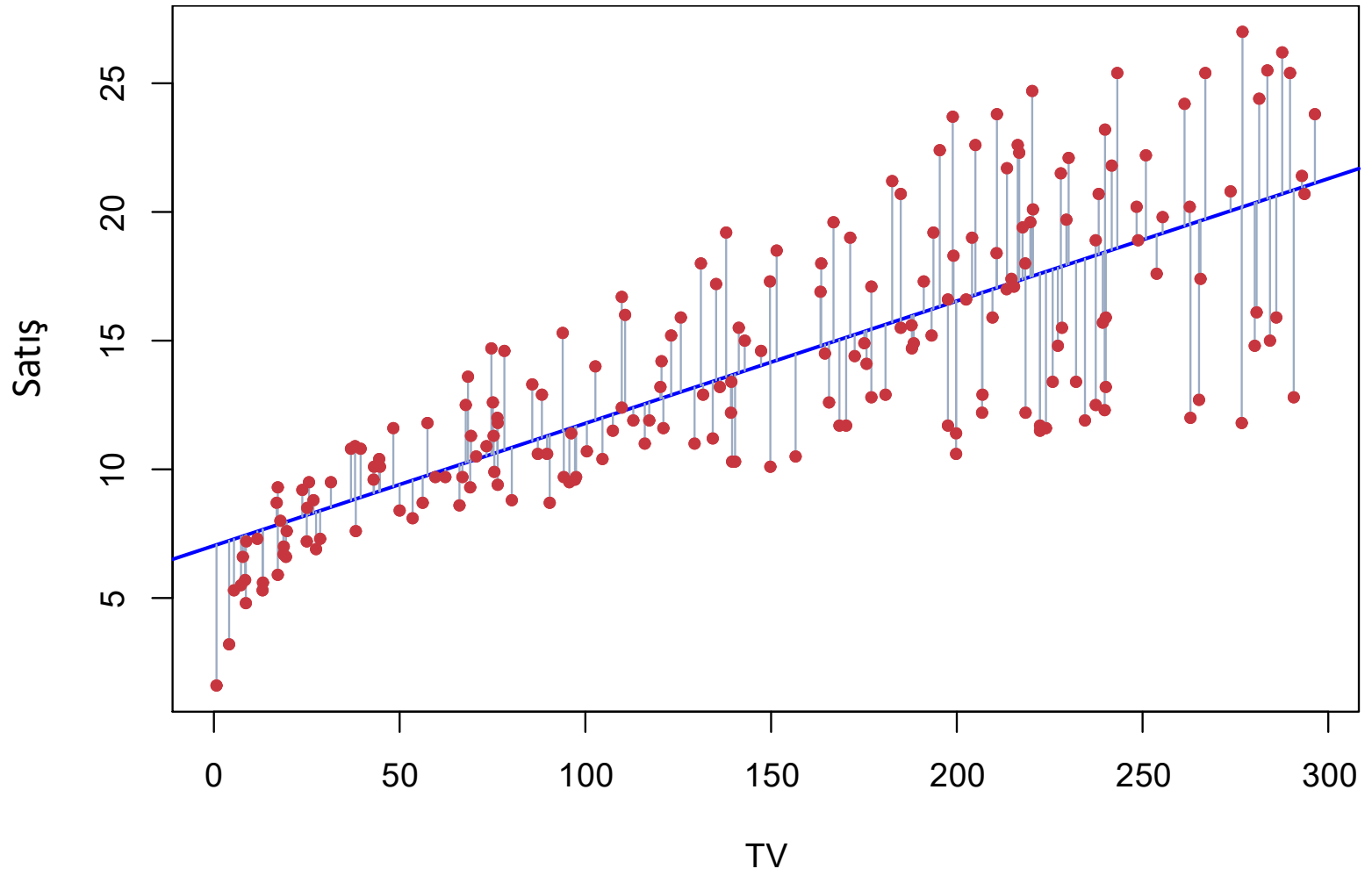
$$y_i \approx \underbrace{\hat{\beta}_0 + \hat{\beta}_1 x_i}_{\hat{y}_i}, \quad i = 1, \dots, n$$

$$y_i \approx \underbrace{\hat{\beta}_0 + \hat{\beta}_1 x_i}_{\hat{y}_i}, \quad i = 1, \dots, n$$

$$e_i = y_i - \hat{y}_i \quad i. \text{ kalıntı (ith residual)}$$

Kalıntı Kareler Toplamı (Residual Sum of Squares)

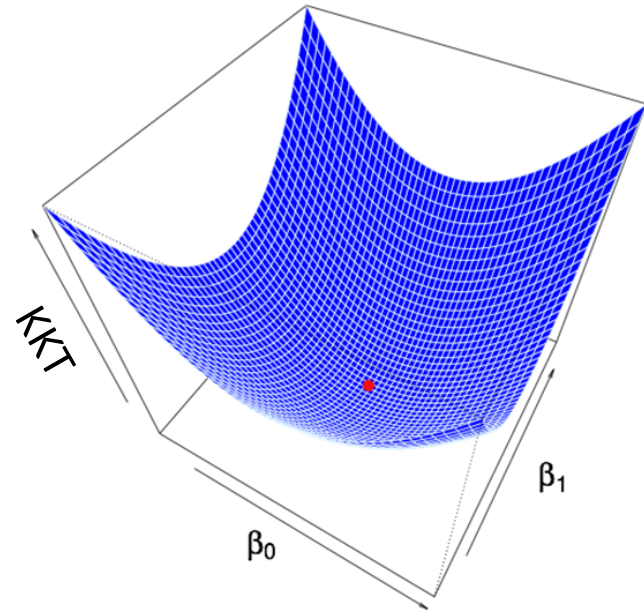
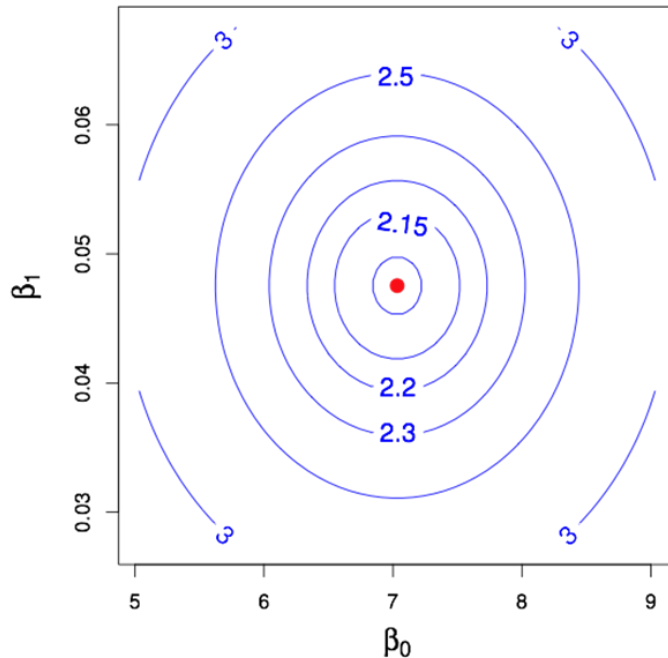
$$\text{KKT} = e_1^2 + e_2^2 + \dots + e_n^2$$



## En Küçük Kareler (EKK) Yöntemi (Least Squares Method)

$$\min_{\beta_0, \beta_1} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

### Dışbükey Fonksiyon



## En Küçük Kareler (EKK) Yöntemi (Least Squares Method)

$$\min_{\beta_0, \beta_1} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$



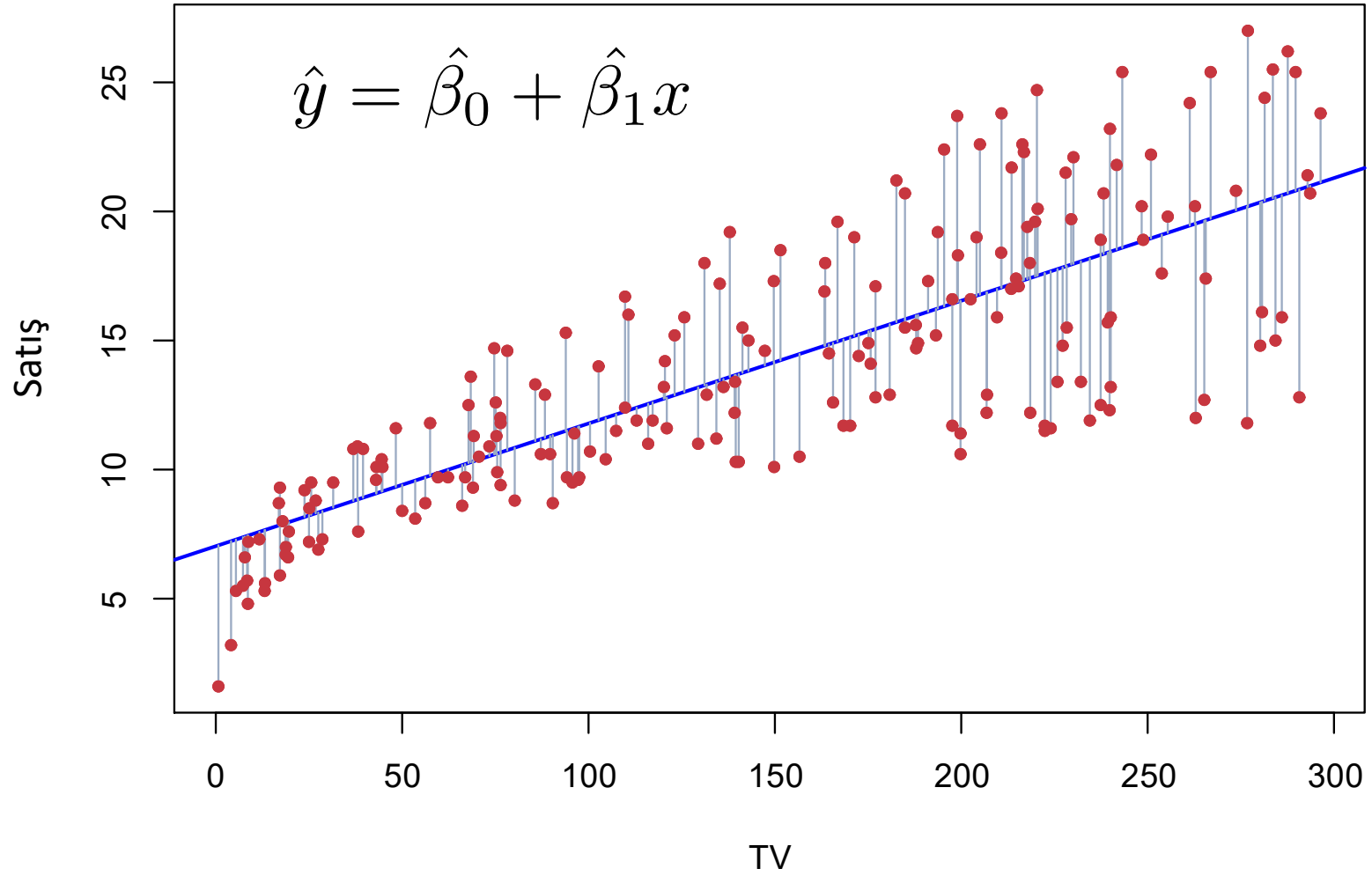
$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$



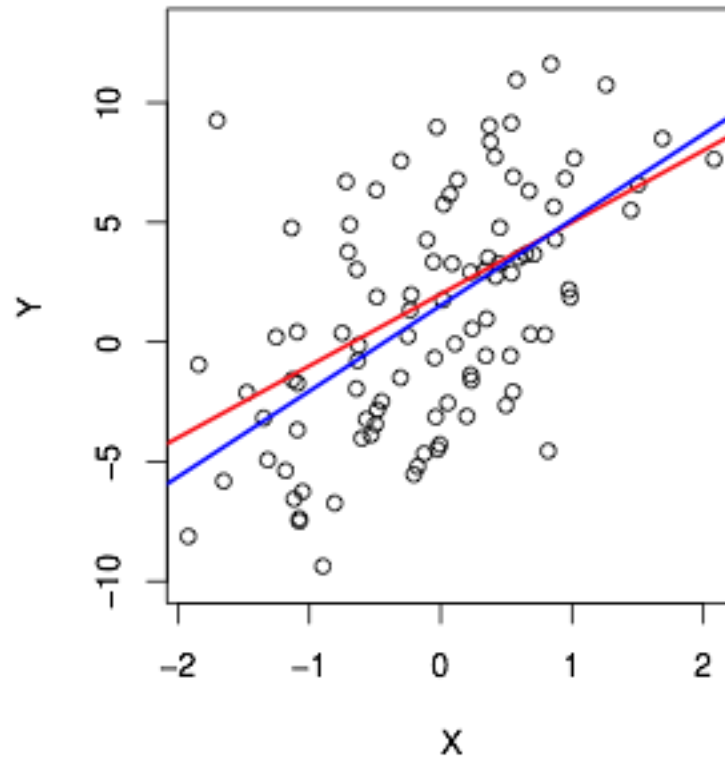


$$Y = f(X) + \epsilon$$

$$Y = \beta_0 + \beta_1 X + \epsilon$$

anakütle bağlanım doğrusu  
(population regression line)

$$Y = 2 + 3X + \epsilon$$

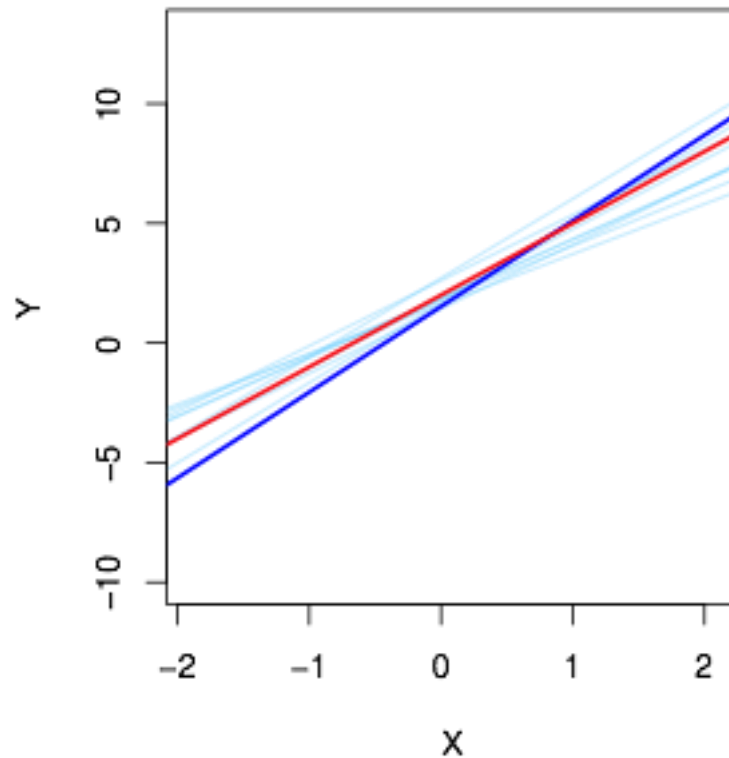


$$\{(x_i^{(1)}, y_i^{(1)}) : i = 1, \dots, n\} \quad \xRightarrow{\text{EKK}} \quad \hat{\beta}_0^{(1)}, \hat{\beta}_1^{(1)}$$

$$\{(x_i^{(2)}, y_i^{(2)}) : i = 1, \dots, n\} \quad \xRightarrow{\text{EKK}} \quad \hat{\beta}_0^{(2)}, \hat{\beta}_1^{(2)}$$

$\vdots$

$$\{(x_i^{(N)}, y_i^{(N)}) : i = 1, \dots, n\} \quad \xRightarrow{\text{EKK}} \quad \hat{\beta}_0^{(N)}, \hat{\beta}_1^{(N)}$$



$$\begin{array}{lll}
 \{(x_i^{(1)}, y_i^{(1)}) : i = 1, \dots, n\} & \xRightarrow{\text{EKK}} & \hat{\beta}_0^{(1)}, \hat{\beta}_1^{(1)} \\
 \{(x_i^{(2)}, y_i^{(2)}) : i = 1, \dots, n\} & \xRightarrow{\text{EKK}} & \hat{\beta}_0^{(2)}, \hat{\beta}_1^{(2)} \\
 \vdots & & \\
 \{(x_i^{(N)}, y_i^{(N)}) : i = 1, \dots, n\} & \xRightarrow{\text{EKK}} & \hat{\beta}_0^{(N)}, \hat{\beta}_1^{(N)}
 \end{array}
 \left. \vphantom{\begin{array}{l} \hat{\beta}_0^{(1)}, \hat{\beta}_1^{(1)} \\ \hat{\beta}_0^{(2)}, \hat{\beta}_1^{(2)} \\ \hat{\beta}_0^{(N)}, \hat{\beta}_1^{(N)} \end{array}} \right\} \begin{array}{l} \text{örneklem} \\ \text{İSTATİSTİK!} \end{array}$$

## Parametreler

SH: Standart Hata (Standard Error)

$$\text{SH}(\hat{\beta}_0)^2 = \sigma^2 \left[ \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]$$

$$\text{SH}(\hat{\beta}_1)^2 = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\sigma^2 = \text{Var}(\epsilon)$$

~ %95 Güven Aralığı (Confidence Interval)

$$\left[ \hat{\beta}_0 - 2 \text{ SH}(\hat{\beta}_0), \hat{\beta}_0 + 2 \text{ SH}(\hat{\beta}_0) \right]$$

$$\left[ \hat{\beta}_1 - 2 \text{ SH}(\hat{\beta}_1), \hat{\beta}_1 + 2 \text{ SH}(\hat{\beta}_1) \right]$$

[6,130; 7,935]

[0,042; 0,053]

~ %95 Güven Aralığı (Confidence Interval)

$$\left[ \hat{\beta}_0 - 2 \text{ SH}(\hat{\beta}_0), \hat{\beta}_0 + 2 \text{ SH}(\hat{\beta}_0) \right]$$

$$\left[ \hat{\beta}_1 - 2 \text{ SH}(\hat{\beta}_1), \hat{\beta}_1 + 2 \text{ SH}(\hat{\beta}_1) \right]$$

Hipotez Testi (Hypothesis Testing)

$$Y = \beta_0 + \beta_1 X + \epsilon$$

$H_0 : \beta_1 = 0$  sıfır hipotezi (null hypothesis)

$H_a : \beta_1 \neq 0$  karşı hipotez (alternative hypothesis)

$t$  testi,  $p$  değeri ...

# Model

KSH: Kalıntı Standart Hata (Residual Standard Error)

$$\text{KSH} = \sqrt{\frac{\text{KKT}}{n - 2}}$$

$R^2$  İstatistiği ( $R^2$  Statistics)

$$R^2 = 1 - \frac{\text{KKT}}{\text{TKT}}$$

$$\text{TKT} = \sum_{i=1}^n (y_i - \bar{y})^2$$

TKT: Tüm Kareler Toplamı

Ölçüm	Değer
KSH	3,26
$R^2$ İstatistiği	0,61

# Çoklu Bağlanım

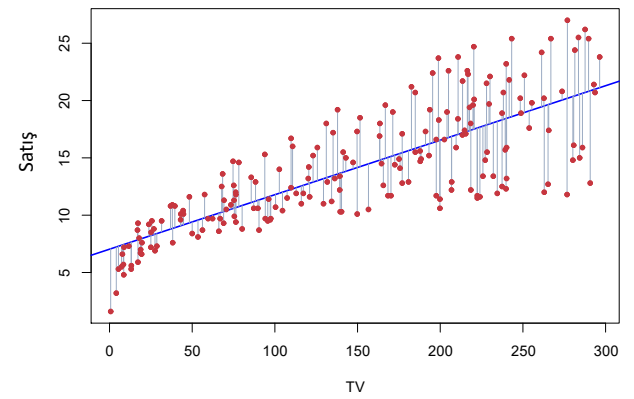
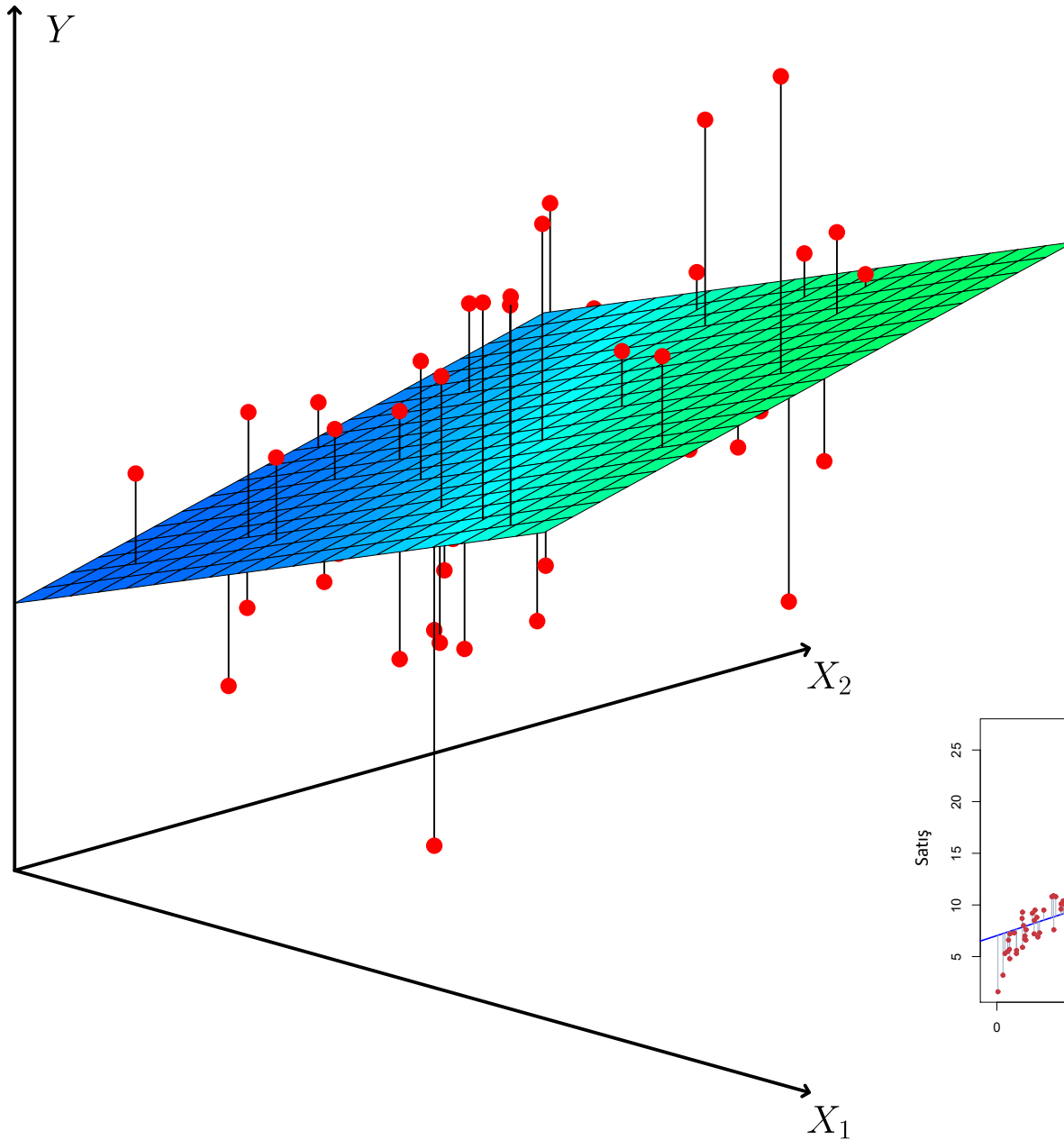
$$Y \approx \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$$

eğitim verisi  
 $\{(x_i, y_i) : 1, \dots, n\}$

$$y_i \approx \underbrace{\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \cdots + \hat{\beta}_p x_{ip}}_{\hat{y}_i}, \quad i = 1, \dots, n$$

$$\text{KKT} = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$





EKK

$$\min_{\beta_0, \beta_1, \dots, \beta_p} \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2$$



Dışbükey Fonksiyon



$$\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$$

$$\min_{\beta_0, \beta_1, \dots, \beta_p} \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_1^\top \\ 1 & x_2^\top \\ \vdots & \vdots \\ 1 & x_n^\top \end{bmatrix}_{n \times (p+1)} \quad \mathbf{y}^\top = (y_1, \dots, y_n)$$

$$\boldsymbol{\beta}^\top = (\beta_0, \beta_1, \dots, \beta_p)$$

$$\hat{\boldsymbol{\beta}}_{\text{LS}} = \arg \min_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

(tam kerte (full rank) varsayımı ile)

$$\hat{\boldsymbol{\beta}}_{\text{LS}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

$$\hat{y}_0 = (1 \ x_0^\top) \hat{\boldsymbol{\beta}}_{\text{LS}}$$

## Basit Bağlanım (satış – gazete)

	<b>Katsayı</b>	<b>St. Hata</b>	<b><i>t</i> testi</b>	<b><i>p</i> değeri</b>
<b>Kesme nok.</b>	12,351	0,621	19,88	< 0,0001
<b>Gazete</b>	0,055	0,017	3,30	0,00115

## Çoklu Bağlanım (satış – TV, radyo, gazete)

	<b>Katsayı</b>	<b>St. Hata</b>	<b><i>t</i> testi</b>	<b><i>p</i> değeri</b>
<b>Kesme nok.</b>	2,939	0,3119	9,42	< 0,0001
<b>TV</b>	0,046	0,0014	32,81	< 0,0001
<b>Radyo</b>	0,189	0,0086	21,89	< 0,0001
<b>Gazete</b>	-0,001	0,0059	-0,18	0,8599

	<b>Katsayı</b>	<b>St. Hata</b>	<b><i>t</i> testi</b>	<b><i>p</i> değeri</b>
<b>Kesme nok.</b>	2,939	0,3119	9,42	< 0,0001
<b>TV</b>	0,046	0,0014	32,81	< 0,0001
<b>Radyo</b>	0,189	0,0086	21,89	< 0,0001
<b>Gazete</b>	-0,001	0,0059	-0,18	<b>0,8599</b>

### İlinti Matrisi

	<b>TV</b>	<b>Radyo</b>	<b>Gazete</b>	<b>Satış</b>
<b>TV</b>	1,0000	0,0548	0,0568	0,7822
<b>Radyo</b>		1,0000	<b>0,3541</b>	0,5762
<b>Gazete</b>			1,0000	0,2283
<b>Satış</b>				1,0000

# Parametreler

## Hipotez Testi

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_a : \text{En az bir } \beta_j \neq 0$$

## $F$ -İstatistiği ( $F$ -Statistics)

$$F = \frac{(\text{TKT} - \text{KKT})/p}{\text{KKT}/(n - p - 1)}$$

Ölçüm	Değer
KSH	1,69
$F$ İstatistiği	570

# Model

KSH

$$\text{KSH} = \sqrt{\frac{\text{KKT}}{n - p - 1}}$$

$R^2$  İstatistiği

$$R^2 = 1 - \frac{\text{KKT}}{\text{TKT}}$$

# Kategorik Değişkenler

İki Seviyeli  
(kadın, erkek)

$$x_i = \begin{cases} 1, & \text{eğer } i. \text{ kişi kadın ise;} \\ 0, & \text{eğer } i. \text{ kişi erkek ise.} \end{cases}$$

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i = \begin{cases} \beta_0 + \beta_1 + \epsilon_i, & \text{eğer } i. \text{ kişi kadın ise;} \\ \beta_0 + \epsilon_i, & \text{eğer } i. \text{ kişi erkek ise.} \end{cases}$$



# Kategorik Değişkenler

## Çok Seviyeli

(esmer, kumral, sarışın)

$$x_{i1} = \begin{cases} 1, & \text{eğer } i. \text{ kişi esmer ise;} \\ 0, & \text{eğer } i. \text{ kişi esmer değilse.} \end{cases}$$

$$x_{i2} = \begin{cases} 1, & \text{eğer } i. \text{ kişi kumral ise;} \\ 0, & \text{eğer } i. \text{ kişi kumral değilse.} \end{cases}$$

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i = \begin{cases} \beta_0 + \beta_1 + \epsilon_i, & \text{eğer } i. \text{ kişi esmer ise;} \\ \beta_0 + \beta_2 + \epsilon_i, & \text{eğer } i. \text{ kişi kumral ise;} \\ \beta_0 + \epsilon_i, & \text{eğer } i. \text{ kişi sarışın ise.} \end{cases}$$

# Dikkat!

- Doğrusal olmayan girdi-çıkı ilişkisi
- İlintili hata terimleri (correlated error terms)
- Hata terimlerinin sabit olmayan varyansı (nonconstant variance)
- Aykırı değerler (outliers) ve *ayrık* değerler
- Doğrudaşlık (collinearity)

# Özet

- Basit Bağlanım
- En Küçük Kareler Yöntemi
- Parametre ve model istatistikleri
- Çoklu Bağlanım
- Kategorik Değişkenler
- Dikkat edilecek birkaç nokta