

# Makine Öğrenmesi

## Boyut Küçültme ve Düzenlileştirme

**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

Öznitelikler  
(Değişkenler)

$$X_1, \cancel{X_2}, X_3, \dots, \cancel{X_{p-1}}, X_p$$

$$Y \approx \beta_0 + \beta_1 X_1 + \underbrace{\cancel{\beta_2}}_0 X_2 + \dots + \underbrace{\cancel{\beta_{p-1}}}_0 X_{p-1} + \beta_p X_p$$

**Tahmin İsabetinin  
Artması**

**Model Yorumunun  
Kolaylaşması**

$n \gg p \longrightarrow$  **EKK**  
(tek çözüm)  $\longrightarrow$  Düşük tahmin varyansı

$p > n \longrightarrow$  **EKK**  
(pek çok çözüm)  $\longrightarrow$  Tahmin varyansı  $\uparrow \infty$

# Alt Küme Seçimi (Subset Selection)

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

$$X_1, X_2, X_3, \dots, X_{p-1}, X_p$$

$$\{\} \xrightarrow{\text{KKT} \downarrow} \mathcal{M}_0 = \{\}$$

$$\{1\}, \{2\}, \dots, \{k\}, \dots, \{p\} \xrightarrow{\text{KKT} \downarrow} \mathcal{M}_1 = \{k\}$$

$$\{1, 2\}, \{1, 3\}, \dots, \{u, v\}, \dots, \{p-1, p\} \xrightarrow{\text{KKT} \downarrow} \mathcal{M}_2 = \{u, v\}$$

$\vdots$

$$\{1, 2, \dots, p\} \xrightarrow{\text{KKT} \downarrow} \mathcal{M}_p = \{1, 2, \dots, p\}$$

$2^p$  alt küme!

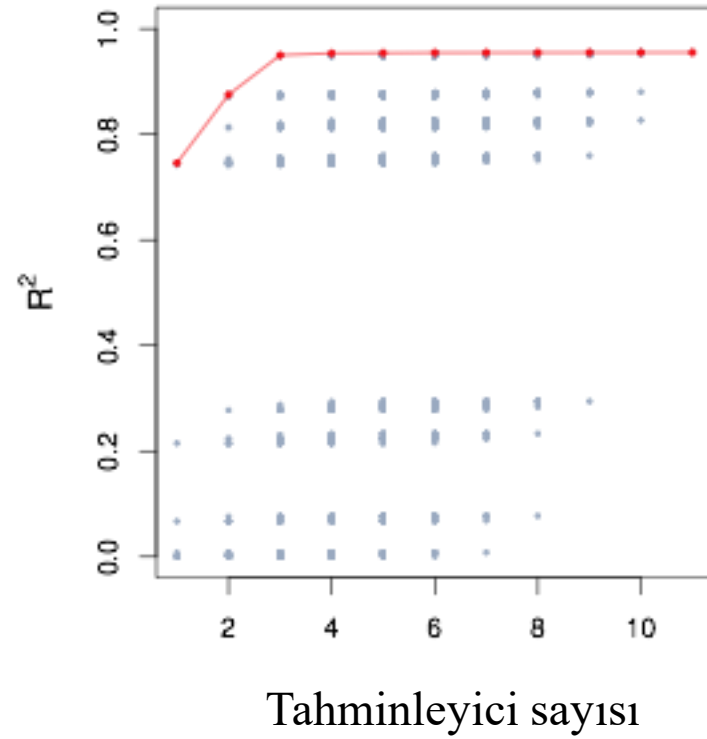
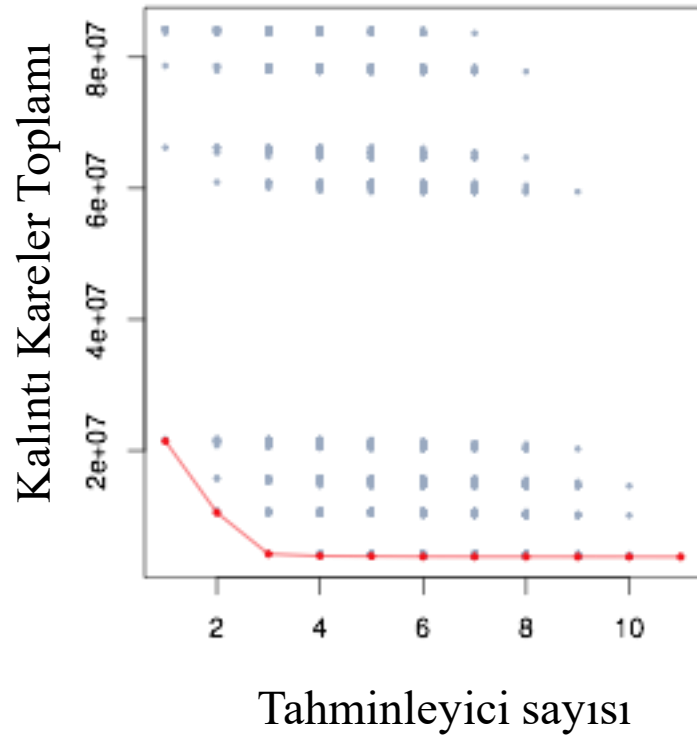
## İleri Adımlı (Forward Stepwise) Alt Küme Seçimi

$$\begin{aligned} \{\} &\xrightarrow{\text{KKT} \downarrow} \mathcal{M}_0 = \{\} \\ \{1\}, \{2\}, \dots, \{\ell\}, \dots, \{p\} &\xrightarrow{\text{KKT} \downarrow} \mathcal{M}_1 = \{\ell\} \\ \{\ell, 2\}, \{\ell, 3\}, \dots, \{\ell, w\}, \dots, \{\ell, p\} &\xrightarrow{\text{KKT} \downarrow} \mathcal{M}_2 = \{\ell, w\} \\ &\vdots \end{aligned}$$

## Geri Adımlı (Backward Stepwise) Alt Küme Seçimi

$$\begin{aligned} \{1, 2, \dots, p\} &\xrightarrow{\text{KKT} \downarrow} \mathcal{M}_p = \{1, 2, \dots, p\} \\ \{2, 3, \dots, p\}, \{1, 3, \dots, p\}, \dots, \{1, 2, \dots, p-1\} &\xrightarrow{\text{KKT} \downarrow} \mathcal{M}_{p-1} = \{1, 3, \dots, p\} \\ &\vdots \end{aligned}$$

$$\frac{p(p+1)}{2} + 1 \text{ alt küme}$$



$$\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_p ?$$

Değişken  
Sayısı



Eğitim  
Hatası



Test  
Hatası



?

... aşırı öğrenme ?

EKK ile doğrusal model öğrenme

$$C_p = \frac{1}{n}(\text{KKT} + 2d\hat{\sigma}^2)$$

(Akaike Information Criterion)

$$\text{AIC} = \frac{1}{n\hat{\sigma}^2}(\text{KKT} + 2d\hat{\sigma}^2)$$

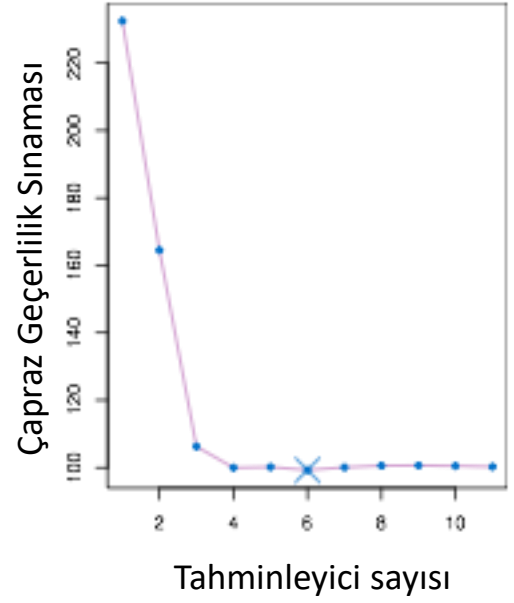
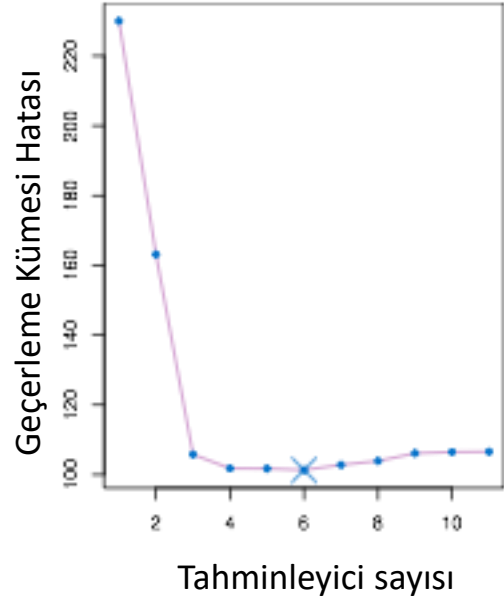
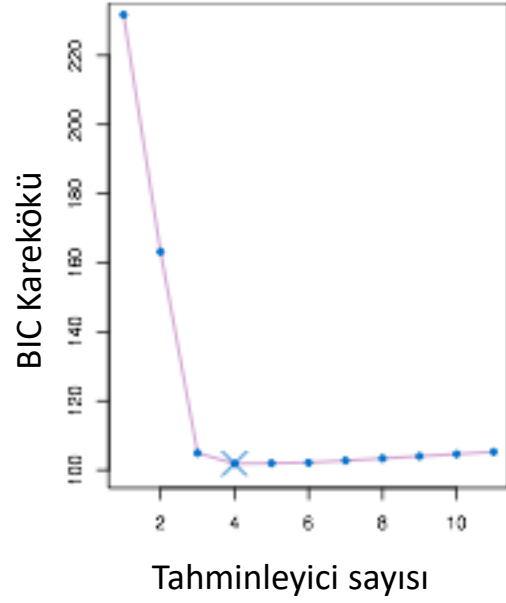
(Bayesian Information Criterion)

$$\text{BIC} = \frac{1}{n\hat{\sigma}^2}(\text{KKT} + \log(n)d\hat{\sigma}^2)$$

$d = 1, 2, \dots, p$

en düşük  
değeri veren  $d \rightarrow \mathcal{M}_d$

Çapraz Geçerlilik Sınaması (Cross Validation)





# Hatırlatma: Doğrusal Bağlanım

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

$$\mathbf{X} = \begin{bmatrix} 1 & x_1^\top \\ 1 & x_2^\top \\ \vdots & \vdots \\ 1 & x_n^\top \end{bmatrix}_{n \times (p+1)} \quad \mathbf{y}^\top = (y_1, \dots, y_n)$$
$$\boldsymbol{\beta}^\top = (\beta_0, \beta_1, \dots, \beta_p)$$

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

Tam kerte (full rank) varsayımı ile

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

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

# Düzenleme (Regularization)

$$Y \approx \beta_0 + \beta_1 X_1 + \underbrace{\beta_2}_{0} X_2 + \cdots + \underbrace{\beta_{p-1}}_{0} X_{p-1} + \beta_p X_p$$

## Çıkıntı Bağlanımı (Ridge Regression)

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

$\lambda$  : ayarlama (tuning) parametresi

## Lasso

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

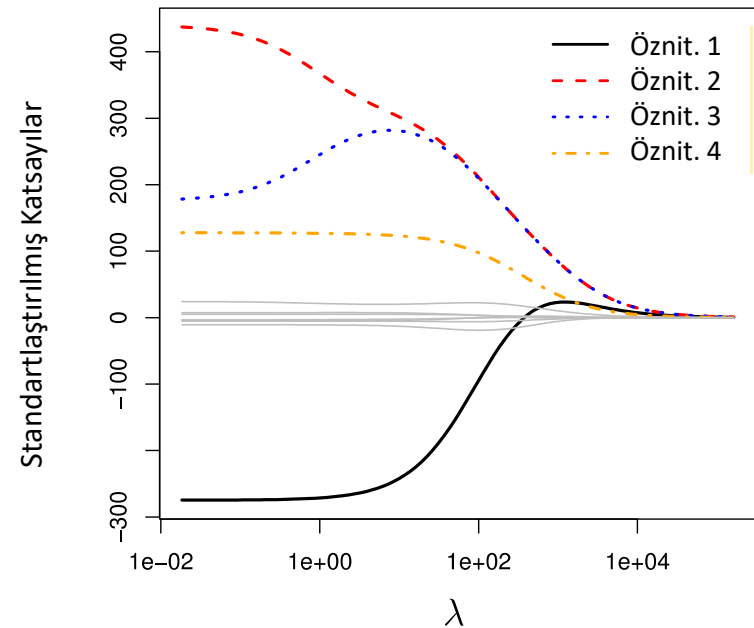
## Çıkıntı Bağlanımı

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

$$\lambda \uparrow \infty \longrightarrow \beta_j \downarrow 0, \quad j = 1, \dots, p$$

### Standartlaştırma

$$\hat{x}_{ij} = \frac{x_{ij}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_{ij})^2}}$$



$$\hat{\beta}_R = \arg \min_{\beta} (\mathbf{y} - \mathbf{X}\beta)^\top (\mathbf{y} - \mathbf{X}\beta) + \lambda \beta^\top \beta$$

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

$$\hat{\beta}_R = (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^\top \mathbf{y}$$

$\beta_0$  hariç  
(ortalanmış girdi)

$$\mathbf{X} = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \vdots \\ x_n^\top \end{bmatrix}_{n \times p}$$

## Tekil Değer Çözüşümü (SVD)

$$\mathbf{X} = \mathbf{U} \mathbf{D} \mathbf{V}^\top$$

$$\mathbf{U} = [\mathbf{u}_1 \ \dots \ \mathbf{u}_p]$$

$$\mathbf{D} = \text{diag}(d_1, \dots, d_p)$$

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

$$= \mathbf{U} \mathbf{U}^\top \mathbf{y}$$

$$= \sum_{j=1}^p \mathbf{u}_j \mathbf{u}_j^\top \mathbf{y}$$

$$\mathbf{X} \hat{\beta}_R = \mathbf{X} (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^\top \mathbf{y}$$

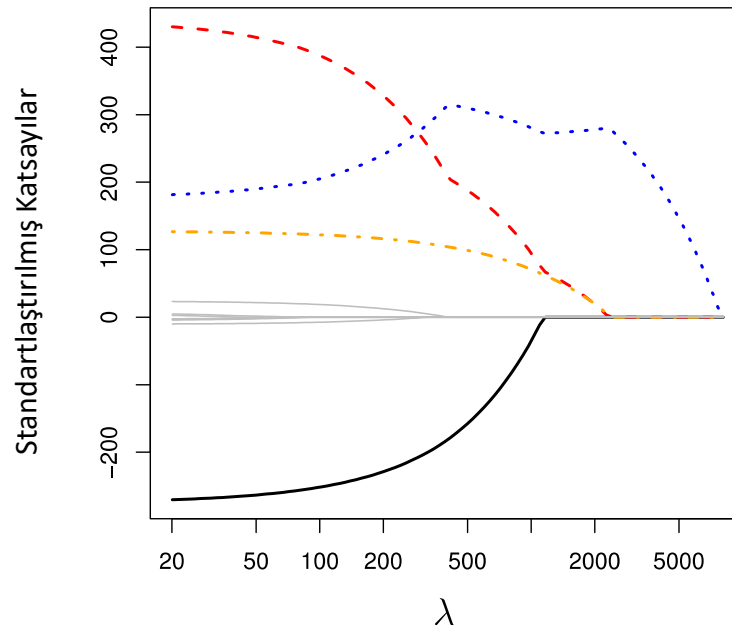
$$= \mathbf{U} \mathbf{D} (\mathbf{D}^2 + \lambda \mathbf{I})^{-1} \mathbf{D} \mathbf{U}^\top \mathbf{y}$$

$$= \sum_{j=1}^p \mathbf{u}_j \frac{d_j^2}{d_j^2 + \lambda} \mathbf{u}_j^\top \mathbf{y}$$

# Lasso

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

$$\lambda \uparrow \infty \longrightarrow \beta_j \downarrow 0, \quad j = 1, \dots, p$$



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

$$\hat{\beta}_{\text{R}} = (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^\top \mathbf{y}$$

$$\hat{\beta}_L = \arg \min_{\beta} (\mathbf{y} - \mathbf{X}\beta)^\top (\mathbf{y} - \mathbf{X}\beta) + \lambda \|\beta\|_1$$



Dışbükey Eniyileme



Analitik çözümü yok

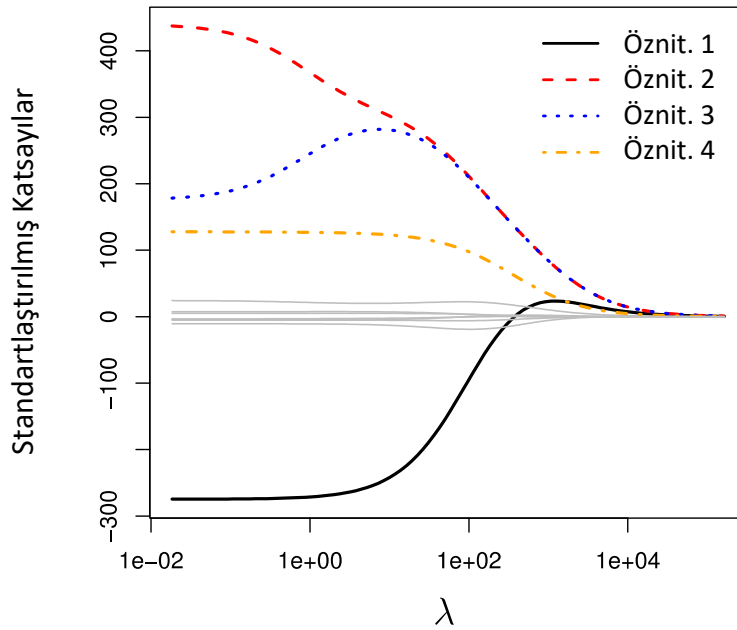
(çok hızlı çözüm yöntemleri mevcut)

$\beta_0$  hariç  
(ortalanmış girdi)

$$\mathbf{X} = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \vdots \\ x_n^\top \end{bmatrix}_{n \times p}$$

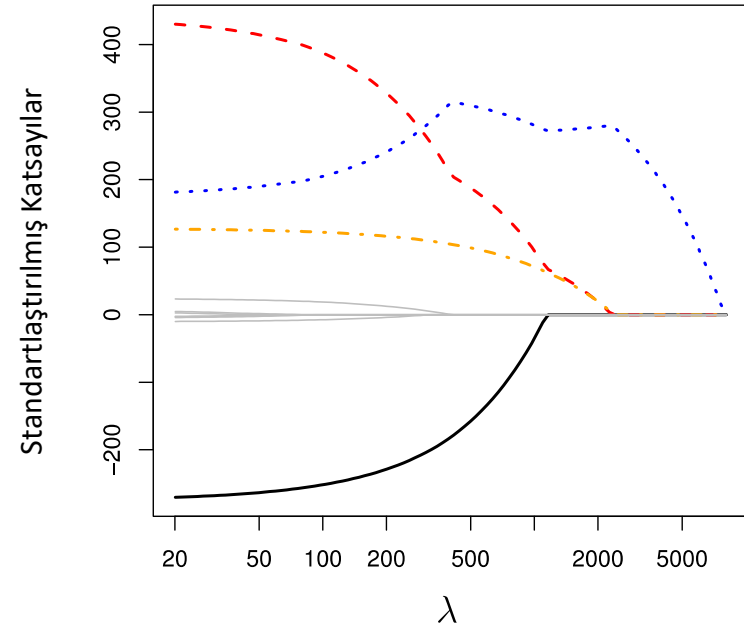
## Çıkıntı Bağlanımı

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



## Lasso

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



## Örnek

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0.01 \\ -0.01 \\ 0.01 \end{bmatrix}$$

$$\sum_{j=1}^3 \beta_j^2 = 3(0.01)^2 = 0.0003$$

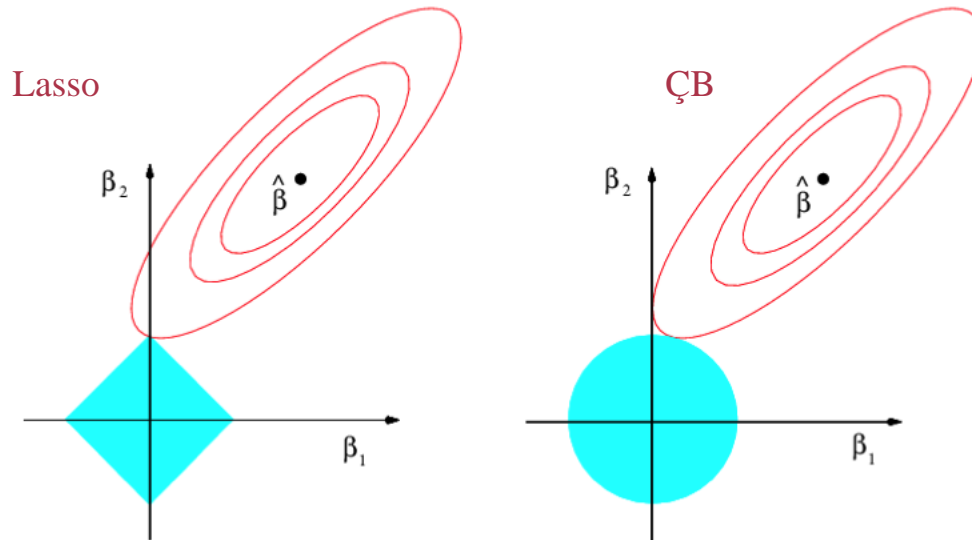
$$\sum_{j=1}^3 |\beta_j| = 3(0.01) = 0.03$$

## Çıkıntı Bağlanımı (ÇB)

$$\min_{\boldsymbol{\beta}} \{ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) : \|\boldsymbol{\beta}\|_2^2 \leq \Delta \}$$

## Lasso

$$\min_{\boldsymbol{\beta}} \{ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) : \|\boldsymbol{\beta}\|_1 \leq \Delta \}$$



En Küçük Kareler Çözümü:  $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$



# Çıkıntı Bağlanımı – Lasso

- Seyrek (bol sıfırlı) çözüm: Lasso
- Çok ve birbirlerine yakın parametreler: Çıkıntı bağlanımı
- Yorumlama kolaylığı: Lasso
- Ayarlama parametresi ( $\lambda$ ) seçimi:

Izgara (grid) arama ve çapraz geçerlilik sınaması

$\lambda$  ↑

Model  
Karmaşıklığı ↓

Varyans ↓

Yanlılık ↑

# Olasılıksal Çıkarsama (Probabilistic Inference)

$y_i \sim N(\hat{y}_i, \sigma^2)$ ,  $i = 1, \dots, n$  ve b.ö.d.

$$\mathbb{P}(\mathbf{y}|\boldsymbol{\beta}) = \prod_{i=1}^n P(y_i|\boldsymbol{\beta})$$

$$\mathbb{P}(\boldsymbol{\beta}|\mathbf{y}) = \frac{\mathbb{P}(\mathbf{y}|\boldsymbol{\beta})\mathbb{P}(\boldsymbol{\beta})}{\mathbb{P}(\mathbf{y})}$$

$$\mathbb{P}(\boldsymbol{\beta}|\mathbf{y}) \propto \mathbb{P}(\mathbf{y}|\boldsymbol{\beta})\mathbb{P}(\boldsymbol{\beta})$$

$$\hat{\boldsymbol{\beta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\beta}} \mathbb{P}(\mathbf{y}|\boldsymbol{\beta})\mathbb{P}(\boldsymbol{\beta})$$

$$= \arg \max_{\boldsymbol{\beta}} \log \mathbb{P}(\mathbf{y}|\boldsymbol{\beta}) + \log \mathbb{P}(\boldsymbol{\beta})$$

$$= \arg \max_{\boldsymbol{\beta}} \sum_{i=1}^n \log \mathbb{P}(y_i|\boldsymbol{\beta}) + \log \mathbb{P}(\boldsymbol{\beta})$$

sabit önsel ,  $\mathbb{P}(\boldsymbol{\beta})$

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{\bullet} &= \arg \max_{\boldsymbol{\beta}} \sum_{i=1}^n \log \mathbb{P}(y_i | \boldsymbol{\beta}) + \log \mathbb{P}(\boldsymbol{\beta}) \\&= \arg \max_{\boldsymbol{\beta}} \sum_{i=1}^n \log \mathbb{P}(y_i | \boldsymbol{\beta}) \\&= \arg \max_{\boldsymbol{\beta}} \sum_{i=1}^n \log \left( \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2}{2\sigma^2}} \right) \\&= \arg \max_{\boldsymbol{\beta}} -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 \\&= \arg \min_{\boldsymbol{\beta}} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 \\&\implies \hat{\boldsymbol{\beta}}_{\bullet} = \hat{\boldsymbol{\beta}}_{\text{LS}}\end{aligned}$$

normal dağılmış önsel ,  $\beta_j \sim N(0, \psi^2)$ ,  $j = 1, \dots, p$  ve b.ö.d.

$$\mathbb{P}(\boldsymbol{\beta}) = \prod_{j=1}^p P(\beta_j)$$

$\beta_0$  hariç  
(ortalanmış veri)

$$\mathbf{X} = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \vdots \\ x_n^\top \end{bmatrix}_{n \times p}$$

$$\hat{\boldsymbol{\beta}}_{\bullet} = \arg \max_{\boldsymbol{\beta}} \sum_{i=1}^n \log \mathbb{P}(y_i | \boldsymbol{\beta}) + \log \mathbb{P}(\boldsymbol{\beta})$$

$$= \arg \max_{\boldsymbol{\beta}} \sum_{i=1}^n \log \left( \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y_i - \sum_{j=1}^p \beta_j x_{ij})^2}{2\sigma^2}} \right) + \sum_{j=1}^p \log \left( \frac{1}{\psi \sqrt{2\pi}} e^{-\frac{\beta_j^2}{2\psi^2}} \right)$$

$$= \arg \min_{\boldsymbol{\beta}} \frac{1}{2\sigma^2} \left( \sum_{i=1}^n (y_i - \sum_{j=1}^p \beta_j x_{ij})^2 + \frac{\sigma^2}{\psi^2} \sum_{j=1}^p \beta_j^2 \right)$$

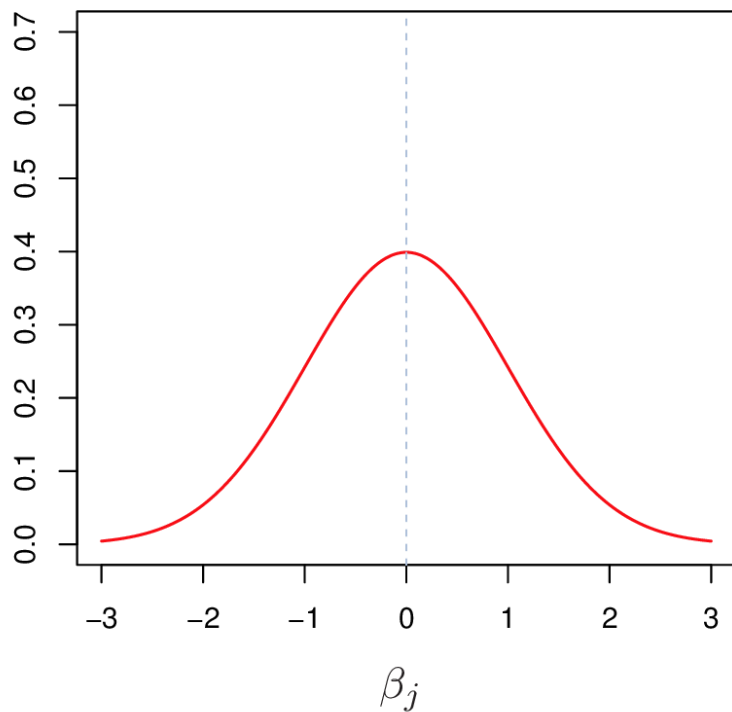
$$= \arg \min_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \frac{\sigma^2}{\psi^2} \boldsymbol{\beta}^\top \boldsymbol{\beta}$$

$$\implies \hat{\boldsymbol{\beta}}_{\bullet} = \hat{\boldsymbol{\beta}}_{\text{R}} \text{ öyle ki } \lambda = \frac{\sigma^2}{\psi^2}$$

Laplace dağılmış önsel  $\beta_j \sim \text{Laplace}(0, \phi)$ ,  $j = 1, \dots, p$  ve b.ö.d.

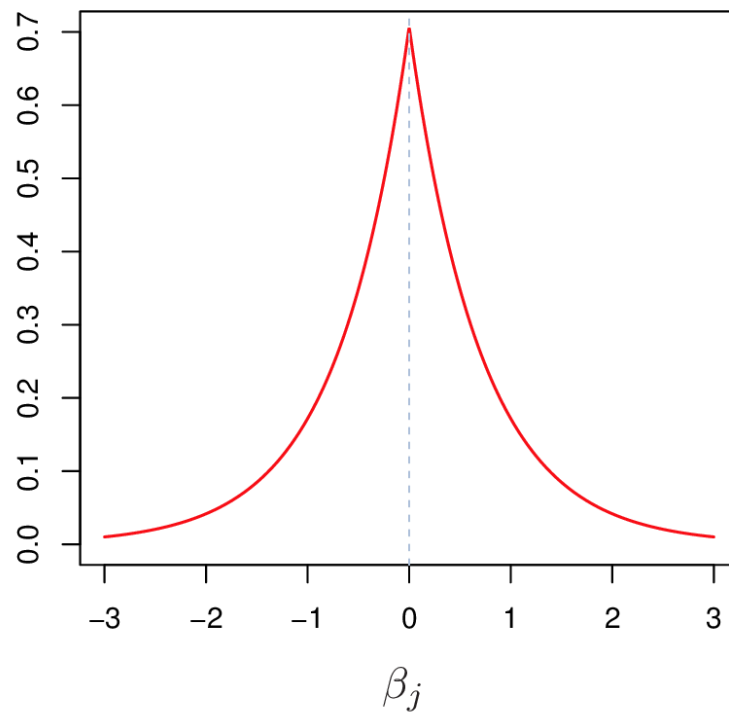
$$\begin{aligned}\hat{\beta}_{\bullet} &= \arg \max_{\beta} \sum_{i=1}^n \log \mathbb{P}(y_i | \beta) + \log \mathbb{P}(\beta) \\ &= \arg \max_{\beta} \sum_{i=1}^n \log \left( \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y_i - \sum_{j=1}^p \beta_j x_{ij})^2}{2\sigma^2}} \right) + \sum_{j=1}^p \log \left( \frac{1}{2\phi} e^{-\frac{|\beta_j|}{2\phi}} \right) \\ &= \arg \min_{\beta} \frac{1}{2\sigma^2} \left( \sum_{i=1}^n (y_i - \sum_{j=1}^p \beta_j x_{ij})^2 + \frac{\sigma^2}{\phi} \sum_{j=1}^p |\beta_j| \right) \\ &= \arg \min_{\beta} (\mathbf{y} - \mathbf{X}\beta)^\top (\mathbf{y} - \mathbf{X}\beta) + \frac{\sigma^2}{\phi} \|\beta\|_1 \\ &\implies \hat{\beta}_{\bullet} = \hat{\beta}_{\text{L}} \text{ öyle ki } \lambda = \frac{\sigma^2}{\phi}\end{aligned}$$

$$\beta_j \sim N(0, \psi^2)$$



$$\hat{\beta}_R = \arg \min_{\beta} (\mathbf{y} - \mathbf{X}\beta)^\top (\mathbf{y} - \mathbf{X}\beta) + \lambda \beta^\top \beta$$

$$\beta_j \sim \text{Laplace}(0, \phi)$$

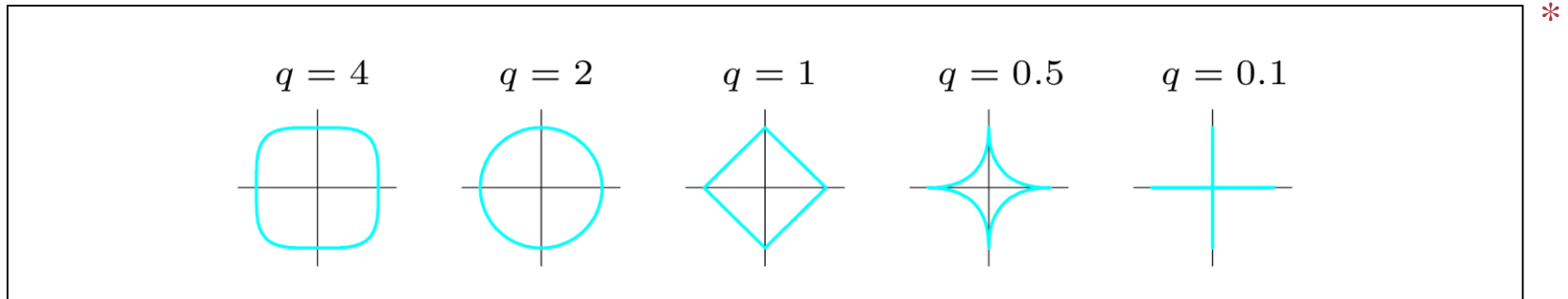


$$\hat{\beta}_L = \arg \min_{\beta} (\mathbf{y} - \mathbf{X}\beta)^\top (\mathbf{y} - \mathbf{X}\beta) + \lambda \|\beta\|_1$$

# Elastic Net

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j|^q$$

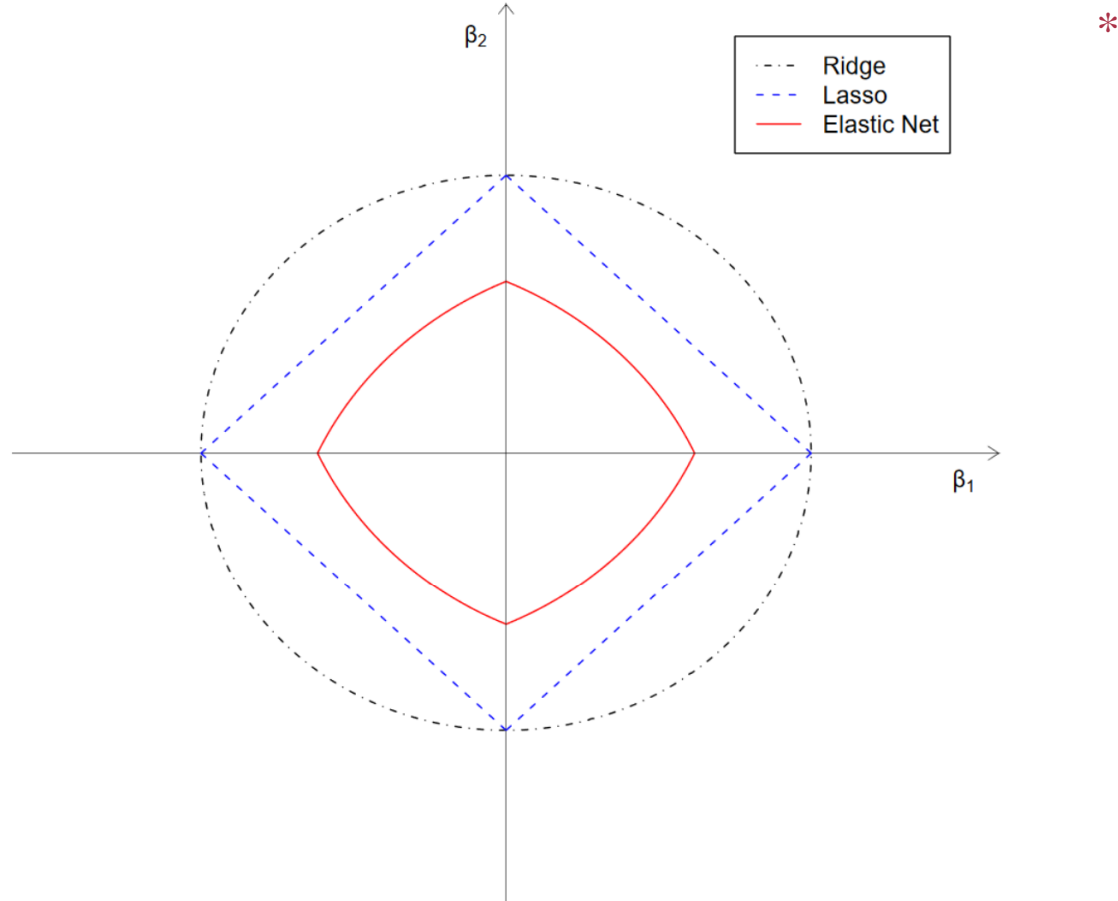
$$q > 0$$



$$\hat{\beta}_{\text{EN}} = \arg \min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p (\alpha \beta_j^2 + (1 - \alpha) |\beta_j|)$$

$$0 \leq \alpha \leq 1$$

## Elastic Net maliyet fonksiyonunun geometrisi





# Pratikte

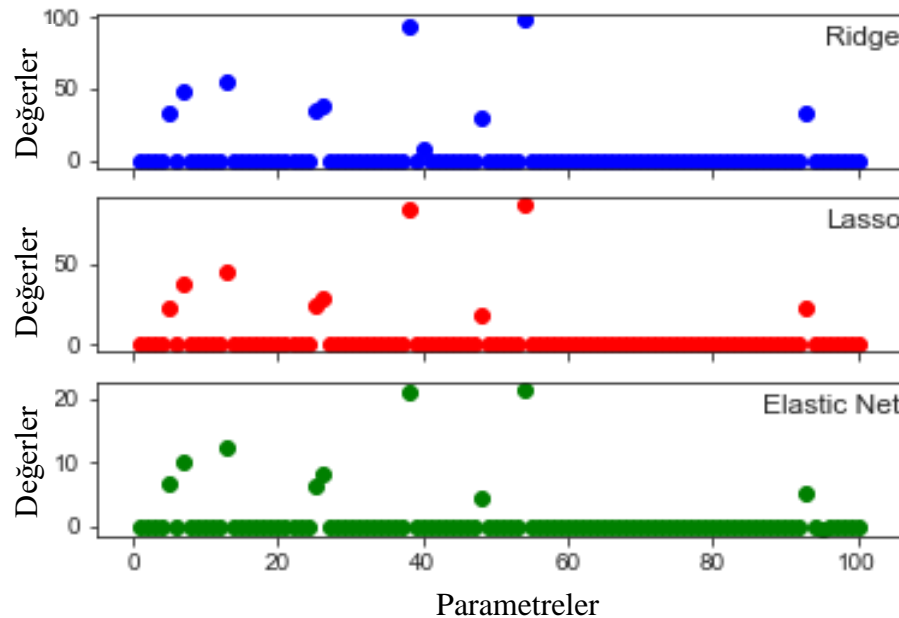
```
1 from sklearn.datasets.samples_generator import make_regression
2 from sklearn.linear_model import Lasso
3 from sklearn.linear_model import Ridge
4 from sklearn.linear_model import ElasticNet
5
6 X, y = make_regression(n_samples=2000, n_features=100, random_state=0)
7
8 rdg = Ridge(alpha=10)
9 lss = Lasso(alpha=10)
10 eln = ElasticNet(alpha=10, l1_ratio=0.7)
11
12 rdg.fit(X, y)
13 lss.fit(X, y)
14 eln.fit(X, y)
```

paketler ve fonksiyonlar

sentetik problem üretimi

üç modelin hazırlanışı

üç modelin uydurumu



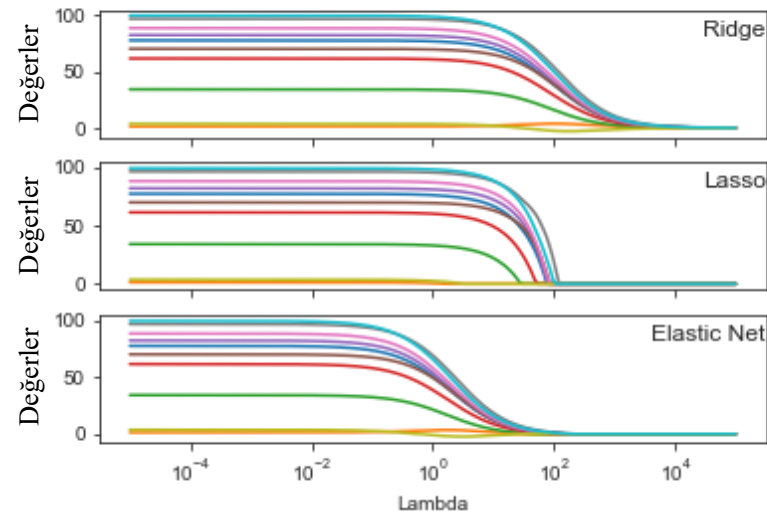
# Pratikte

```
1 import numpy as np
2 from sklearn.datasets.samples_generator import make_regression
3 from sklearn.linear_model import Lasso
4 from sklearn.linear_model import Ridge
5 from sklearn.linear_model import ElasticNet
6
7 X, y = make_regression(n_samples=100, n_features=10, random_state=0, bias=3.5)
8
9 rdg = Ridge()
10 lss = Lasso()
11 eln = ElasticNet()
12
13 rdg_coefs = []; lss_coefs = []; eln_coefs = [];
14 alphas = np.logspace(-5, 5, 200)
15 for a in alphas:
16     rdg.set_params(alpha=a)
17     lss.set_params(alpha=a)
18     eln.set_params(alpha=a)
19     rdg.fit(X, y); lss.fit(X, y); eln.fit(X, y);
20     rdg_coefs.append(rdg.coef_)
21     lss_coefs.append(lss.coef_)
22     eln_coefs.append(eln.coef_)
```

paketler ve fonksiyonlar

üç modelin hazırlanışı

değişken maliyet fonksiyonu  
parametreleri



# Boyut Küçültme

$$X_1, X_2, \dots, X_p$$

$$Z_m = \sum_{j=1}^p \phi_{jm} X_j, \quad m = 1, 2, \dots, M \quad (M < p)$$

$$Y \approx \theta_0 + \theta_1 Z_1 + \theta_2 Z_2 + \dots + \theta_M Z_M$$

Örnek

$$p = 4, M = 2$$

$$Z_1 = \phi_{11} X_1 + \phi_{21} X_2 + \phi_{31} X_3 + \phi_{41} X_4$$

$$Z_2 = \phi_{12} X_1 + \phi_{22} X_2 + \phi_{32} X_3 + \phi_{42} X_4$$

$$Y \approx \theta_0 + \theta_1 Z_1 + \theta_2 Z_2$$

# İzdüşüm (Projection)

$$X_1, X_2, \dots, X_p$$

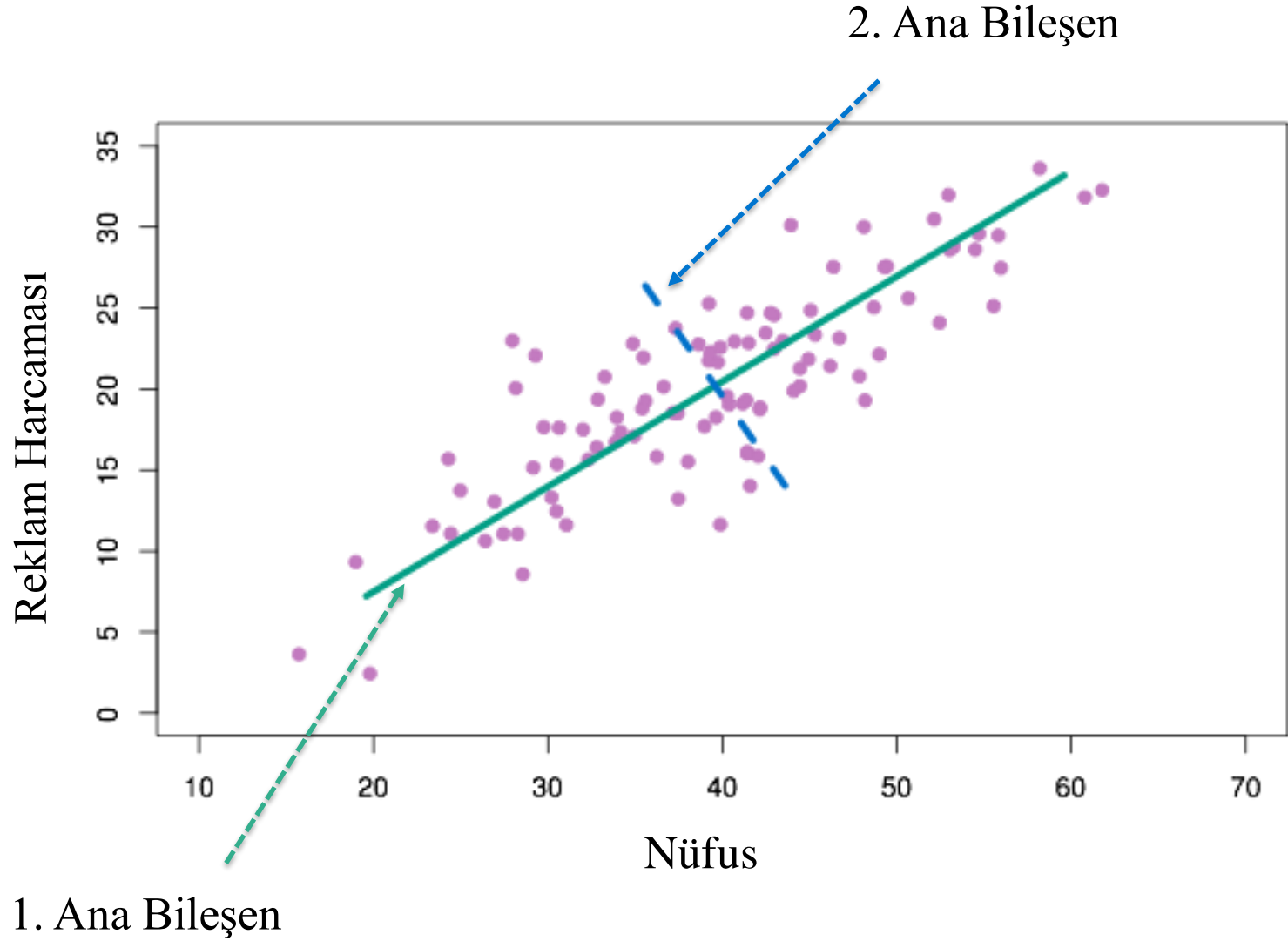


$$Z_m = \sum_{j=1}^p \phi_{jm} X_j, \quad m = 1, 2, \dots, M$$

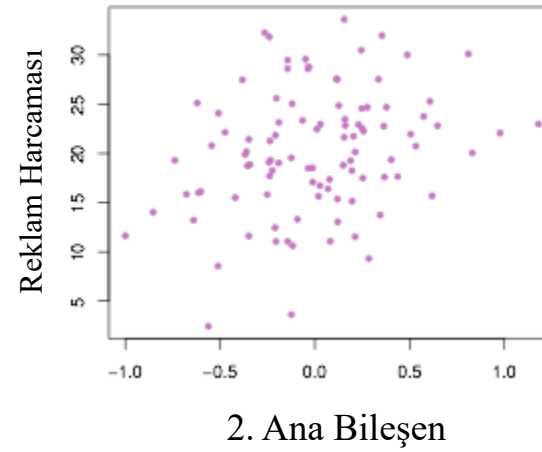
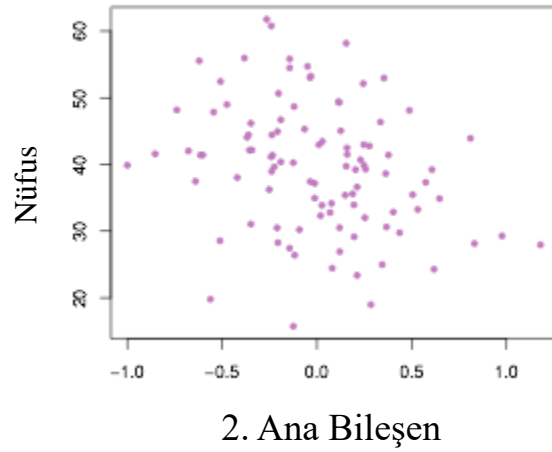
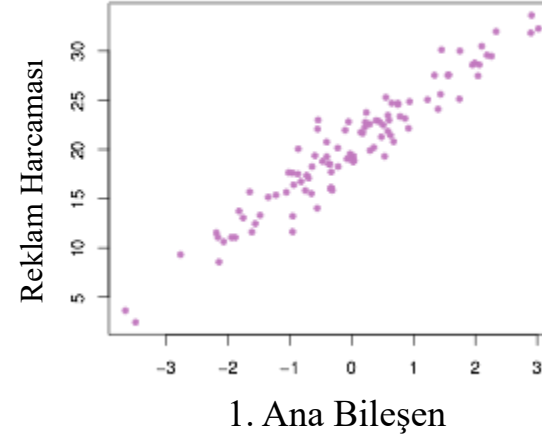
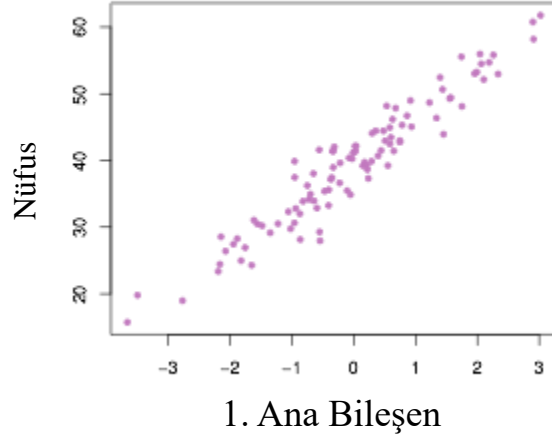
Ana Bileşenler Analizi – ABA  
(Principal Component Analysis)

**Öznitelik seçmek için kullanılmaz!**

# Görsel olarak Ana Bileşenler Analizi



# Görsel olarak Ana Bileşenler Analizi



# Ana Bileşenler Bağlanımı (ABB)

ABA Uygulaması



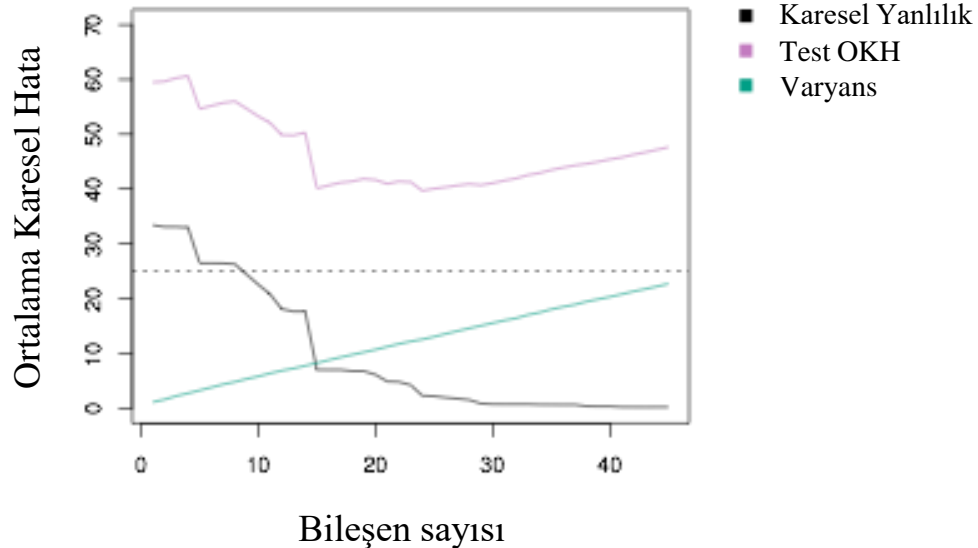
$$Y \approx \theta_0 + \theta_1 Z_1 + \theta_2 Z_2 + \cdots + \theta_M Z_M$$



EKK Uygulaması

Standartlaştırma  
uygulaması (çıkıntı  
bağlanımı gibi)

$M$  seçimi için çapraz  
geçerlilik sınaması  
uygulanabilir



# Pratikte

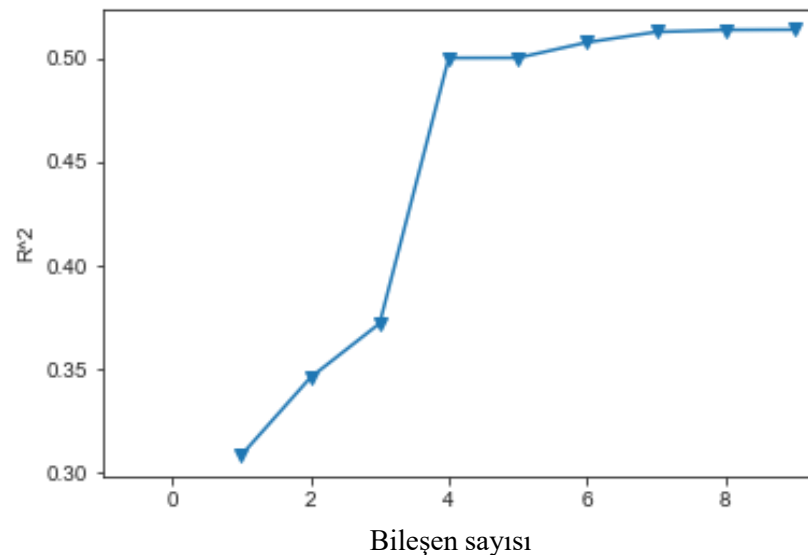
```
1 import numpy as np
2 from sklearn.datasets import load_diabetes
3 from sklearn.decomposition import PCA
4 from sklearn.linear_model import LinearRegression
5 from sklearn.preprocessing import scale
6
7 X, y = load_diabetes(return_X_y=True)
8
9 pca = PCA()
10
11 Z = pca.fit_transform(scale(X))
12
13 regr = LinearRegression()
14 r2scores = []
15
16 p = len(Z[1,:])
17 # Fit to all components one-by-one
18 for i in np.arange(1, p):
19     score = regr.fit(Z[:, :i], y).score(Z[:, :i], y)
20     r2scores.append(score)
```

paketler ve fonksiyonlar

*diabetes* verisinin yüklenmesi

ABA (ölçekleme ile)

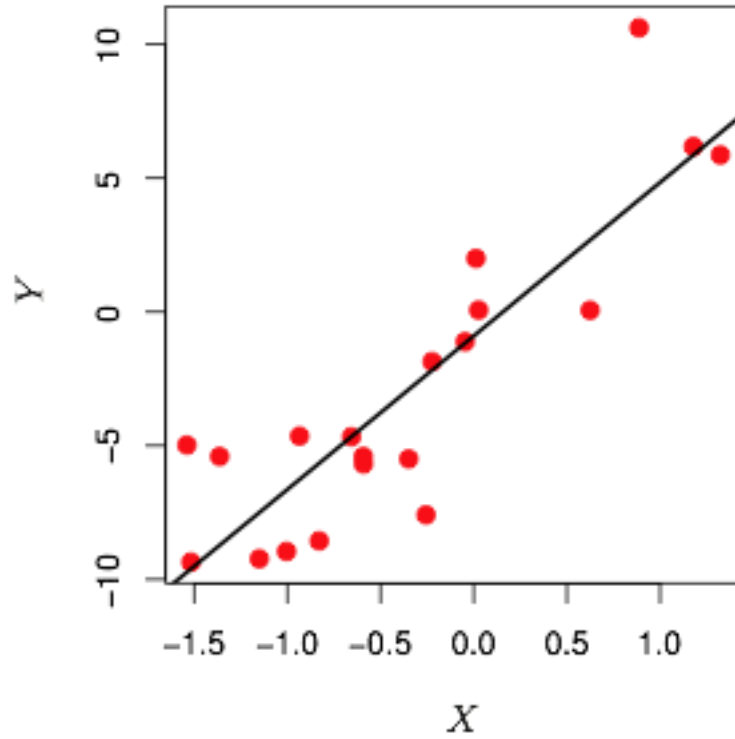
Farklı ana bileşen sayısına göre eğitim hataları



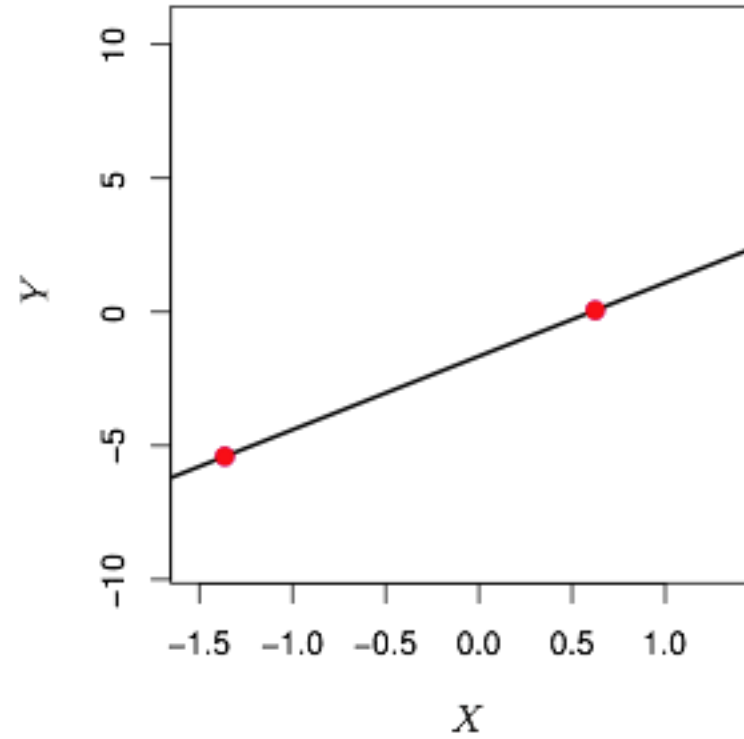


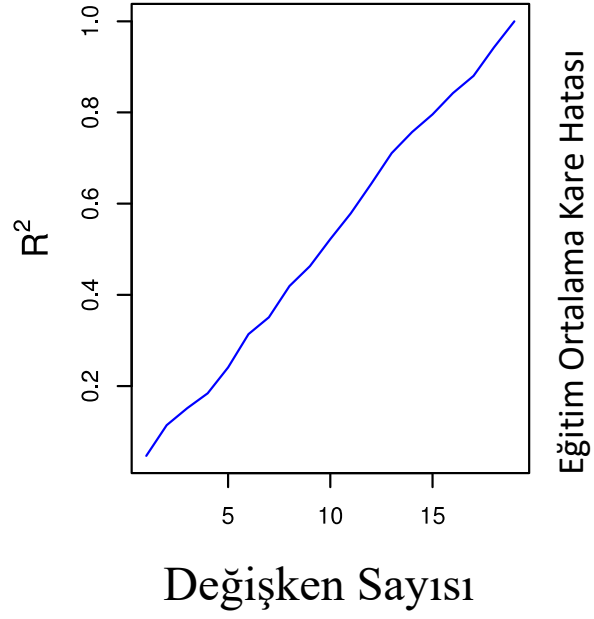
# Yüksek Boyutlar

$$p \ll n$$

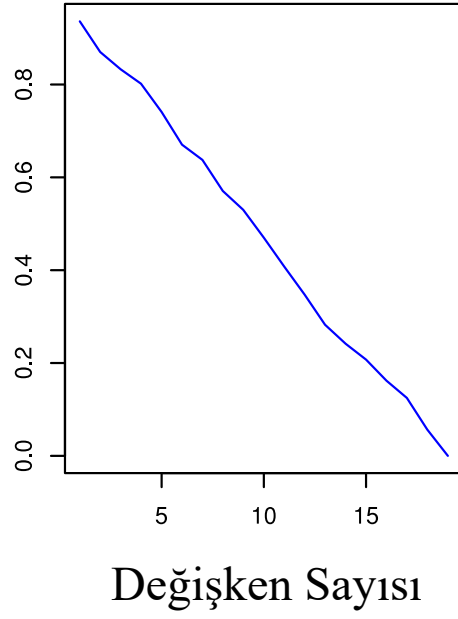


$$p > n$$

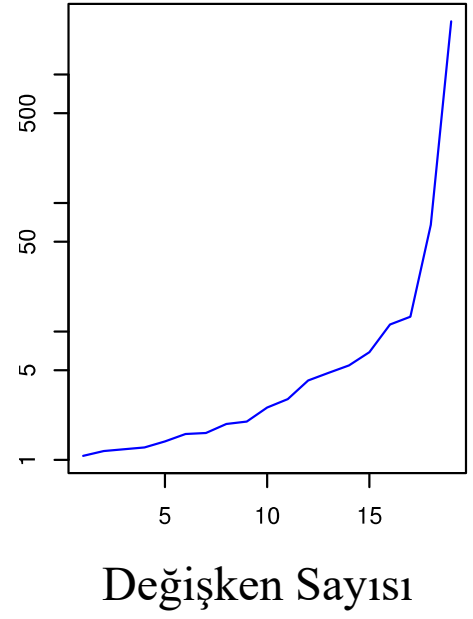




Eğitim Ortalama Kare Hatası



Test Ortalama Kare Hatası



# Dikkat!

$$p > n$$

- Güvenilmez EKK sonuçları
- ~~$R^2$ ,  $p$  değeri, KKT~~
- Çapraz Geçerlilik Sınaması uygulaması
- Alt küme seçiminde güçlük
- Pek çok değişkenin doğrudaş (collinear) olması
- İleri adımlı seçim, çıkıntı bağlantımı, lasso, elastic net veya ABB

# Özet

- Alt Küme Seçimi
  - İleri Adımlı Alt Küme Seçimi
  - Geri Adımlı Alt Küme Seçimi
- Düzenlileştirme
  - Çıkıntı Bağlanımı
  - Lasso
  - Elastic Net
- Boyut küçültme: Ana Bileşenler Bağlanımı
- Yüksek Boyutlar