KOCAELİ ÜNİVERSİTESİ
BİLGİSAYAR MÜHENDİSLİĞİ

Bilgisayar Mühendisliğinde Matematik Uygulamaları

8. Hafta

Yrd. Doç. Dr. A. Burak İNNER

Kocaeli Üniversitesi Bilgisayar Mühendisliği
Yapay Zeka ve Benzetim Sistemleri Ar-Ge Lab.
http://yapbenzet.kocaeli.edu.tr

REGRESYON VE İNTERPOLASYON

Regresyon

- En Küçük Kareler Yöntemi
  - Doğru Uydurma
  - Polinom Uydurma
  - Üstel Fonksiyonlara Eğri Uydurma

İnterpolasyon

- Lagrange İnterpolasyonu (Polinom İnterpolasyon)
- Newton İnterpolasyonu

REGRESYON VE İNTERPOLASYON

Eğri uydurma için, veri hatalarına bağlı olarak birbirinden ayrılan iki yaklaşım vardır:

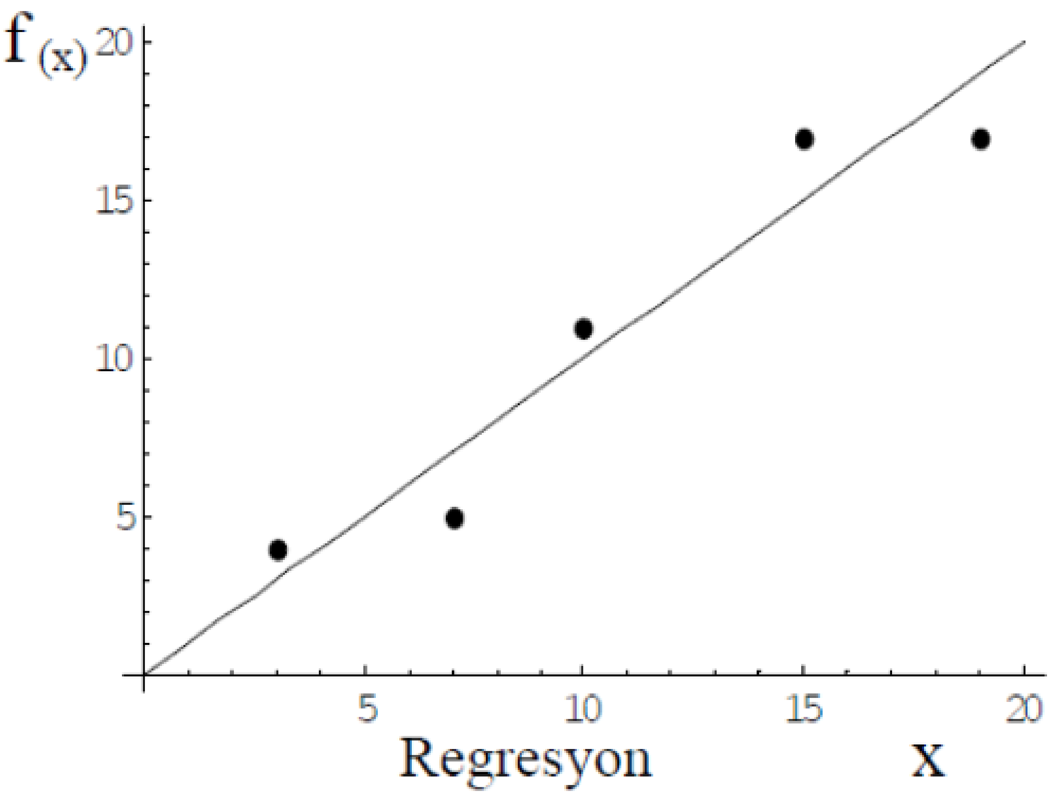
Regresyon

Hata oranı büyük değerlerin, her bir veri noktasından geçmeyen,verilerin genel eğiliminin tek bir eğri ile gösterimidir.(Deneysel veriler için)

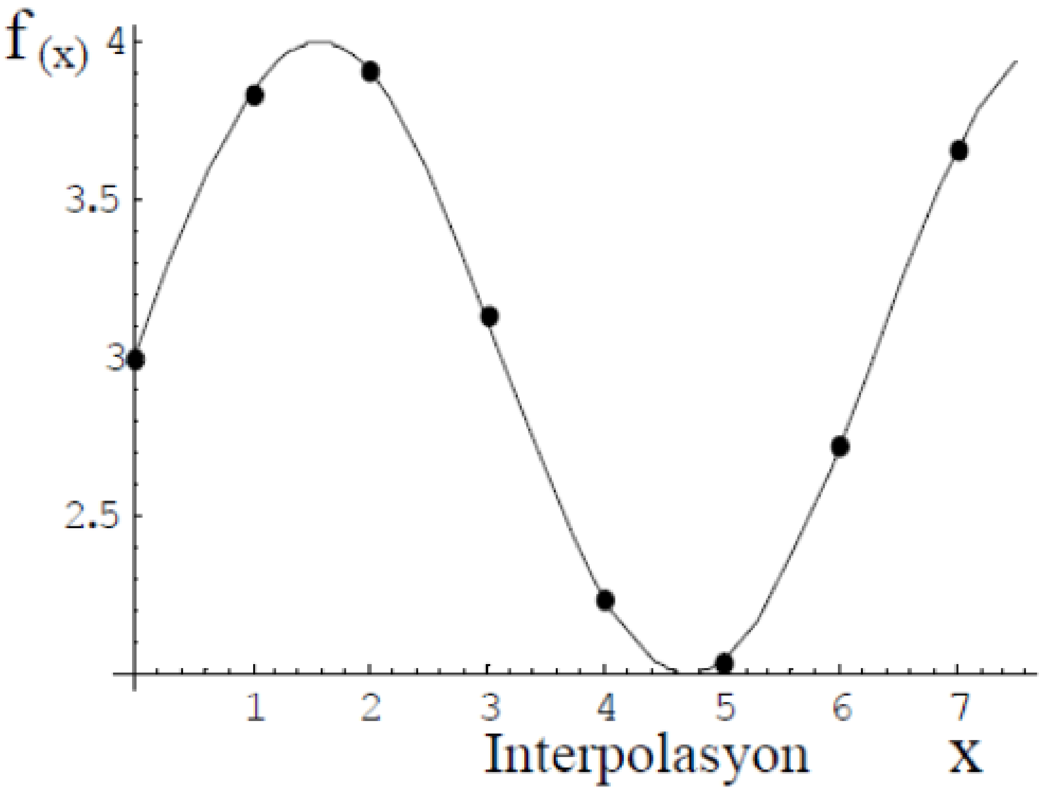
İnterpolasyon

Hata oranı küçük değerlerin, iyi bilinen ayrık noktaların her birinden geçecek şekilde eğri veya eğri uydurularak gösterimidir.

REGRESYON VE İNTERPOLASYON



Regresyon
(yaklaştırma)



Interpolasyon
(ara değeri bulma)

REGRESYON

- Regresyon analizi yaparken en çok kullanılan yöntemlerden biri en küçük kareler yöntemidir.

- En Küçük Kareler Yöntemi gerçek yaşamın çeşitli alanlarında herhangi bir uygulama ile toplanan veriler tablo şekline getirilerek incelenir ve toplanan veriyi modelleyen bir fonksiyon bulunmaya çalışılır.

REGRESYON

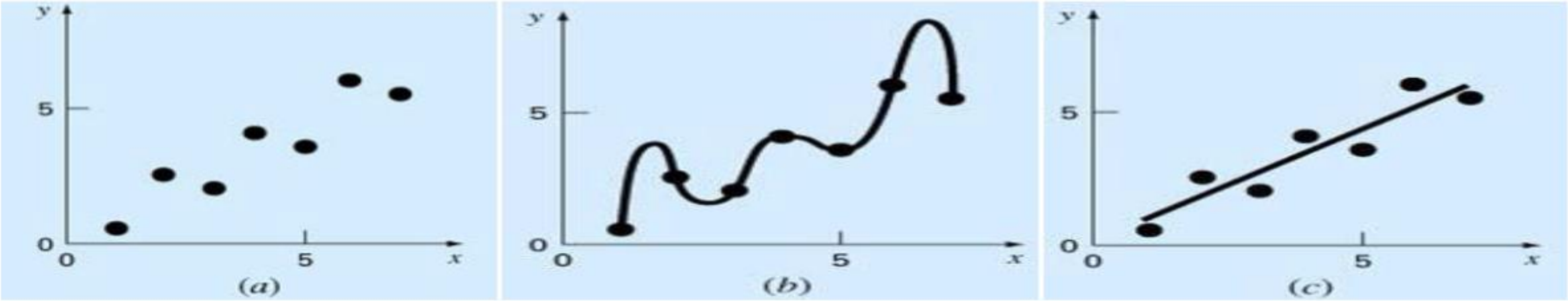
- Çoğu zaman bu veri tablosuna tam olarak uyan bir fonksiyon bulmak mümkün olmaz.
- Veri tablosuna en iyi uyan fonksiyon belirlenmeye çalışılır. Bir veri tablosuna en iyi uyan fonksiyonu bulma sürecine regresyon analizi denir.

REGRESYON

Verilerde önemli hatalar olduğunda, interpolasyon uygun değildir ve ara değerleri tahmin etmek için kullanıldığında tatmin edici sonuçlar vermez.Genellikle deneysel veriler bu tiptedir ve en küçük kareler yöntemi ile gösterimi daha iyi sonuçlar verir.

REGRESYON

✓Verilerinin belirli bir doğruyla gösterilebilmesi durumunda doğru uydurma yöntemine başvurulur.



Veriler　　　　İnterpolasyon　　　　En küçük kareler yöntemi

REGRESYON

- Bu yöntemde doğruya yaklaşımdaki hataların karelerinin toplamını minimum yapacak doğru denklemi araştırılır.

$$y = a_0 + a_1 x + \varepsilon \quad \longrightarrow \quad \varepsilon = y_{gerçek} - y_{yaklaşık} = y - (a_0 + a_1 x)$$

$$S_r = \sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (y - a_0 - a_1 x_i)^2$$

a_0 : kesme noktası

a_1 : eğim

ε : hata veya artık

S_r : hataların kareleri toplamı

REGRESYON

- Hataların kareleri toplamını minimum yapacak a_0 ve a_1 değerleri, bu değerlerin türevlerinin sıfıra eşitlenmesiyle bulunur.Son durumda a_0 ve a_1 aşağıdaki gibi olur.

$$(n)a_0 + (\sum x_i)a_1 = \sum y_i$$

$$(\sum x_i)a_0 + (\sum x_i^2)a_1 = \sum x_i y_i$$

REGRESYON

- Aşağıdaki denklem takımı çözülerek a_0 ve a_1 değerleri bulunur.

$$a_1 = \bar{y} \ - \ a_1\bar{x} = \frac{\sum y_i}{n} - a_1 \frac{\sum x_i}{n}$$

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum \mathrm{x}_i{}^2 - (\sum x_i)^2}$$

Not : $\sum x_i y_i$ ile $\sum x_i \sum y_i$ ve $\sum \mathrm{x}_i{}^2$ *ile* $(\sum x_i)^2$ birbirinden farklı ifadelerdir.

REGRESYON

- Eğri uydurmanın uyumluluğunu belirlemek için korelasyon katsayısı (r) ile belirlenir. Doğrusal regresyon için kolerasyon katsayısı aşağıdaki eşitlik ile hesaplanır:

$$r = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i{}^2 - (\sum x_i)^2}\ \sqrt{n\sum y_i{}^2 - (\sum y_i)^2}}$$

REGRESYON

Tablodaki değerler için düz bir doğru uydurun.

x_i	y_i
1	0.5
2	2.5
3	2.0
4	4.0
5	3.5
6	6.0
7	5 5

REGRESYON

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum \mathrm{x}_i{}^2 - (\sum x_i)^2} = \frac{7\times 119.5 - 28\times 24}{7\times 140 - 28^2} = 0.839$$

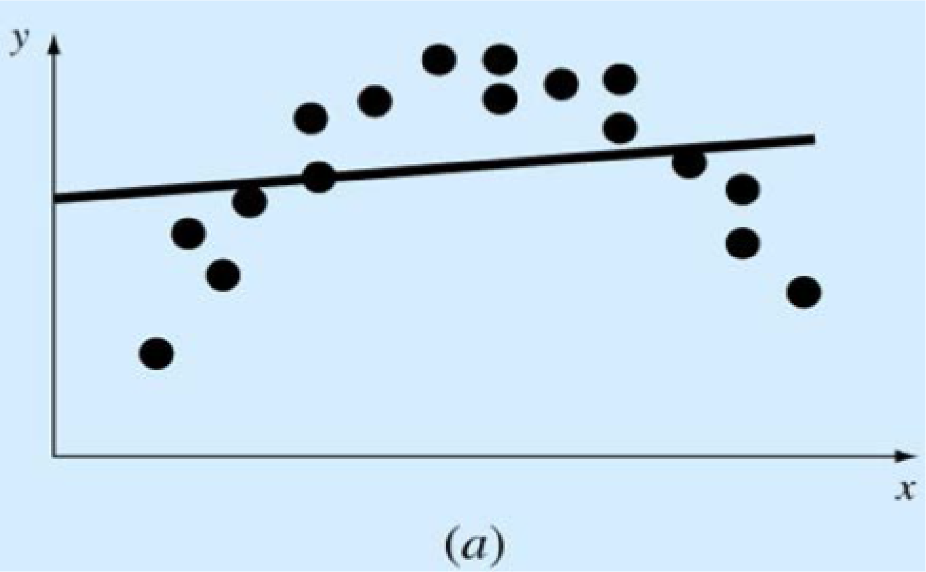
$$a_0 = \bar{y} - a_1\bar{x} = 3.429 - 0.839 \times 4 = 0.073$$

y = A + Bx = 0.073 + 0.839x

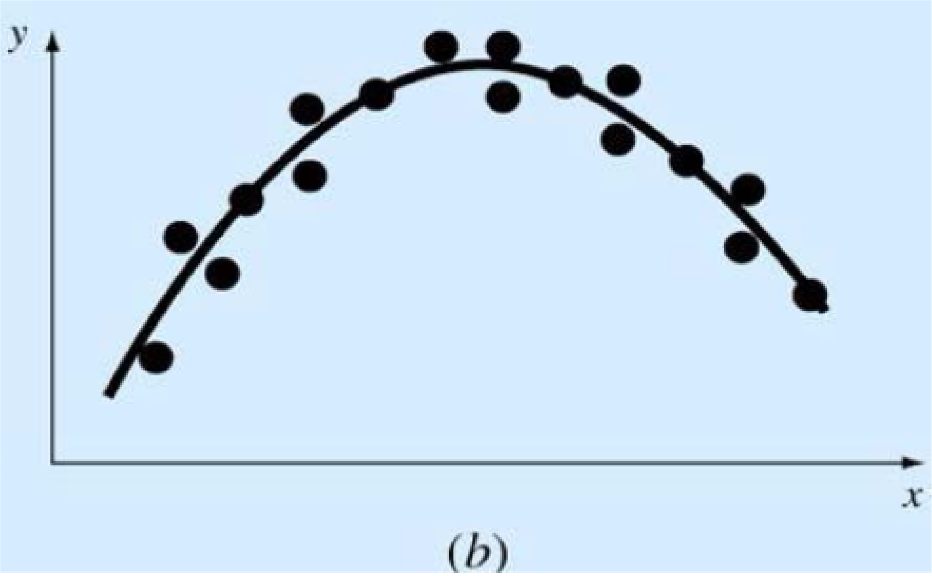
$$r = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum \mathrm{x}_i{}^2 - (\sum x_i)^2}\sqrt{n\sum \mathrm{y}_i{}^2 - (\sum y_i)^2}} = \frac{7\times 119.5 - 28\times 24}{\sqrt{7\times 140 - 28^2} - \sqrt{7\times 105 - 24^2}} = \frac{164.5}{176.53} = 0.932$$

REGRESYON

✓Bazı mühendislik verilerinin belirli bir şekli olsa da düz bir doğruyla gösterilmesi mümkün değildir. Böyle durumlarda bir eğri, verilere daha uyumlu olabilir ve polinom uydurmak doğru daha doğru sonuçlar verir.



Doğrusal regresyona uymayan veriler



Parabolik eğrinin kullanımı

REGRESYON

- En küçük kareler yöntemi, yüksek dereceli polinomlara eğri uydurmak için kolayca genişletilebilir:

$$y = a_0 + a_1 x + a_2 x^2 + \varepsilon$$

$$\rightarrow \varepsilon = y_{gerçek} - y_{yaklaşık} = y - (a_0 + a_1 x + a_2 x^2)$$

$$S_r = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (a_0 + a_1 x + a_2 x^2)^2$$

REGRESYON

- Hatanın karelerinin toplamını minimum yapacak $a_0, a_{1\ ve}\ a_2$ değerleri, bu değerlerin türevlerinin sıfıra eşitlenmesiyle bulunur.Son durumda aşağıdaki eşitlikler elde edilir.

$$n(a_0) + (\sum x_i)\, a_1 + (\sum x_i{}^2)\, a_2 = \sum y_i$$

$$(\sum x_i)\, a_0 + (\sum x_i{}^2)\, a_1 + (\sum x_i{}^3)\, a_2 = \sum x_i y_i$$

$$(\sum x_i{}^2)\, a_0 + (\sum x_i{}^3)\, a_1 + (\sum x_i{}^4)\, a_2 = \sum x_i{}^2 y_i$$

REGRESYON

- Tablodaki değerleri 2. dereceden polinoma yaklaştırın

x_i	y_i
0	2.1
1	7.7
2	13.6
3	27.2
4	40.09
5	61.1

REGRESYON

$$n(a_0) + \left(\sum x_i\right) a_1 + \left(\sum x_i^2\right) a_2 = \sum y_i \rightarrow 6a_0 + 15a_1 + 55a_2 = 152.6$$

$$\left(\sum x_i\right) a_0 + \left(\sum x_i^2\right) a_1 + \left(\sum x_i^3\right) a_2 = \sum x_i y_i \rightarrow 15a_0 + 55a_1 + 255a_2 = 585.6$$

$$\left(\sum x_i^2\right) a_0 + \left(\sum x_i^3\right) a_1 + \left(\sum x_i^4\right) a_2 = \sum x_i^2 y_i \rightarrow 55a_0 + 225a_1 + 979a_2 = 2488.8$$

REGRESYON

- Cramer yöntemiyle çözüm uygulanırsa;

1. A = $\begin{vmatrix} 152.6 & 15 & 55 \\ 585.6 & 55 & 225 \\ 2488.8 & 225 & 979 \end{vmatrix}$ B = $\begin{vmatrix} 6 & 152.6 & 55 \\ 15 & 585.6 & 225 \\ 55 & 2488.8 & 979 \end{vmatrix}$ C = $\begin{vmatrix} 6 & 15 & 152.6 \\ 15 & 55 & 585.6 \\ 55 & 225 & 2488.8 \end{vmatrix}$ D = $\begin{vmatrix} 6 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{vmatrix}$

2. $a_0 = \frac{detA}{detD} = \frac{979}{3920} = 2.479$

$a_1 = \frac{detB}{detD} = \frac{9248.4}{3920} = 2.359 \quad \rightarrow \; a_0 + a_1 x + a_2 x^2 = 2.479 + 2.359x + 1.861x^2$

$a_2 = \frac{detC}{detD} = \frac{7294}{3920} = 1.861$

REGRESYON

y = $a_0 x^{a_1}$ şeklindeki lineer olmayan denklemin, kuvvet fonksiyonunu elde etmek için her iki tarafın doğal logaritması alınır:

In y = In($a_0 x^{a_1}$) $\rightarrow$ In y = In(a_0) + a_1In(x)

In y = Y In a_0 = A a_1=B In x = X alınırsa Y = A+BX

ε = $\sum_{i=1}^{n}(Y_i - A - Bx_i)^2$