Pertemuan IV

PENCOCOKAN KURVA:

ANALISA REGRESI

Materi Minggu Ini

- Pencocokan Kurva
- Regresi Kuadrat Terkecil (RKT)
  - RKT untuk Kurva Linier
  - RKT untuk Kurva Non-Linier
- Regresi Polynomial
- Tugas IV

Background Layout Theme Transition

Pencocokan Kurva (1)

Seringkali data tersajikan dalam bentuk rangkaian nilai diskrit (deretan angka² dalam urutan yang kontinu), tanpa disertai bentuk fungsi yang menghasilkan data tsb.

Dalam kasus di atas, kita dapat men-"generate" fungsi sederhana untuk mengaproksimasi bentuk fungsi sebenarnya dengan memanfaatkan rangkaian data yang ada.

Mengapa bentuk fungsi begitu penting bagi kita?...

Selain untuk memenuhi kebutuhan proses numeris, seperti integrasi atau mendapatkan solusi pendekatan dari persamaan differensial, seringkali kita harus menganalisa tren atau melakukan pengujian hipotesa terhadap nilai² diskrit yang dihasilkan oleh fungsi tsb.

Pencocokan Kurva (2)

Terhadap keberadaan rangkaian/pasangan data yang tidak diketahui fungsi asalnya, terdapat 2 hal yang dapat dilakukan :

Pertama, berusaha mencari bentuk kurva (fungsi) yang mewakili rangkaian data diskrit tersebut.

Kedua, berusaha meng-estimasi/memperkirakan nilai fungsi pada titik^2 tertentu yang belum diketahui.

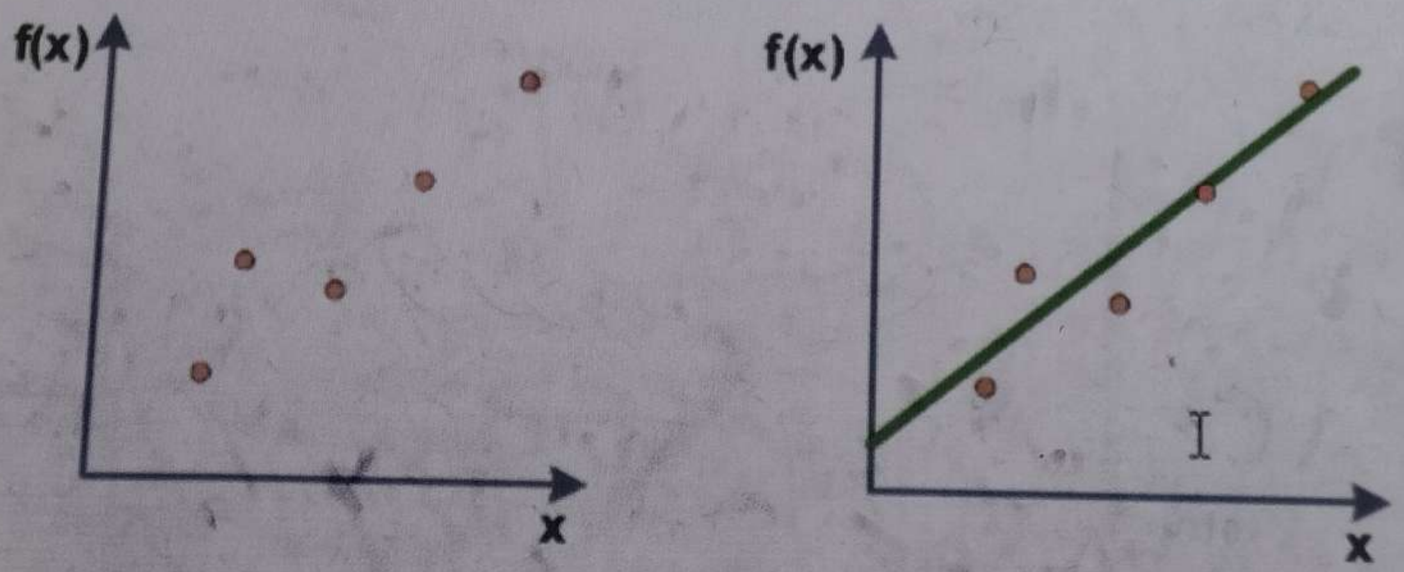
Keduanya dikenal sebagai teknik Pencocokan Kurva (*curve fitting*).

Pencocokan Kurva (3)

Pendekatan^2 yang lazim digunakan untuk melakukan Pencocokan Kurva antara lain adalah :

Regresi Kuadrat Terkecil (*least-square regresion*)

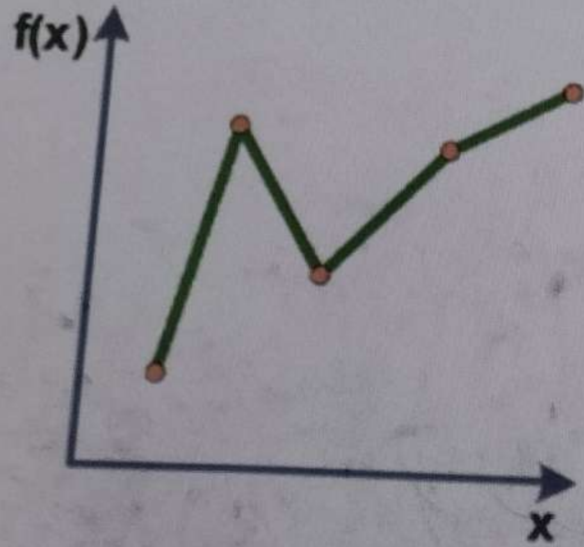
Metode ini digunakan apabila data yang tersaji memiliki tingkat kesalahan berarti (akurasi) rendah. Anda hanya perlu membuat sebuah garis lurus yang merepresentasikan tren dari data^2 tsb secara umum.

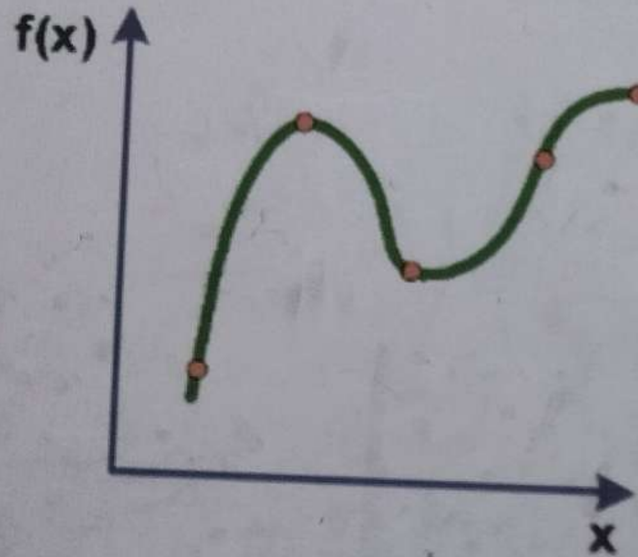


Pencocokan Kurva (3)

Interpolasi

Jika tingkat ketelitian data yang kita miliki lebih baik, maka metode *Interpolasi* dapat dipakai. Kita gunakan segmen² garis lurus untuk menghubungkan titik² data (interpolasi linier). Atau, dengan menggunakan kurva (interpolasi polynomial) untuk memperoleh hasil yang lebih baik.





Analisis regresi menggunakan sedikit notasi dan perhitungan statistik. Ini artinya, ada sedikit yang perlu anda ingat kembali...

Sedikit Notasi Statistik

Tahun (x_i)	Debit Air = y_i (m^3/det)
1991	8,52
1992	3,33
1993	7,85
1994	7,65
1995	10,91
1996	4,17
1997	3,40
1998	8,00
1999	13,4
2000	5,40
2001	8,87
2002	4,73
2003	7,40
2004	6,88
2005	5,00

Rerata/rata-rata data ($\overline{y}$) adalah jumlah nilai data (Σy)dibagi jumlah data (n) :

$$\overline{y} = \frac{\Sigma y_i}{n}$$

Deviasi Standar (σ) atau penyebaran nilai data adalah :

$$\sigma = \sqrt{\frac{\Sigma (y_i - \overline{y})^2}{n - 1}}$$

Atau dapat pula dinyatakan dalam bentuk Varians (σ^2) :

$$\sigma^2 = \frac{\Sigma (y_i - \overline{y})^2}{n - 1}$$

Regresi Kuadrat Terkecil (1)

Jika data yang tersaji memiliki tingkat kesalahan yang cukup signifikan, maka penggunaan interpolasi bukanlah pilihan yang bijaksana. Karena (kemungkinan besar) hasil pendekatannya akan kurang memuaskan.

Ada cara yang lebih mudah, yaitu dengan membuat kurva yang dapat mewakili titik² data tersebut.

Katakan kurva ini adalah kurva dari fungsi g(x) yang 'mirip' dengan fungsi sebenarnya.

Tetapi jika penyebaran titik datanya sangat besar? ...

Rasanya koq sulit cara di atas bisa berhasil dengan baik ☹

Kecuali jika, ... kita dapat membuat kurva buatan, g(x), yang mampu meminimalkan perbedaan (selisih) dengan kurva aslinya, f(x).

Teknik ini yang disebut dengan metode Regresi Kuadrat Terkecil.

Regresi Kuadrat Terkecil (2)

Secara umum prosedur untuk mengaplikasikan metode Regresi Kuadrat Terkecil (RKT) ini adalah sbb :

1. Titik² data diplot ke dalam koordinat cartesian. Dari pola datanya bisa dilihat, apakah kurva pendekatan yang akan kita buat berbentuk garis lurus atau garis lengkung;
2. Tentukanlah sebuah fungsi g(x) yang dpt mewakili fungsi titik² data f(x) :
$$g(x) = a_0 + a_1x + a_2x^2 + ... + a_rx^r$$
3. Jika a_0, a_1, ..., a_r adalah parameter fungsi g(x), maka tentukan nilai parameter² tsb sdmk hingga g(x) dpt mendekati titik² data;
4. Jika koordinat titik² data adalah $M(x_i, y_i)$, maka selisih dengan fungsi g(x) adalah :
$$\begin{aligned} {}_iE &= M_i - G_i \\ &= y_i - g(x_i; a_0, a_1, ..., a_r) \\ &= y_i - (a_0 + a_1x_i + a_2x_i^2 + a_3x_i^3 + ... + a_rx_i^r) \end{aligned}$$

Regresi Kuadrat Terkecil (3)

5. Tingkat kesalahan fungsi $g(x)$ diukur melalui rumus berikut :

$$D^2 = \sum_{i=1}^{n} E_i^2 = \sum_{i=1}^{n} \{ y_i - g(x_i)\}^2$$

6. Cari nilai parameter a_0, a_1, ..., a_r sdmk hingga D^2 dapat seminimum mungkin. D^2 bernilai minimum jika turunan pertamanya terhadap a_0, a_1, ..., a_r = 0.

$$\frac{\partial D^2}{\partial a_0} = 0 \quad \frac{\partial D^2}{\partial a_1} = 0 \quad \ldots \quad \frac{\partial D^2}{\partial a_r} = 0$$

7. Persamaan no 6 di atas akan memberikan nilai parameter a_0, a_1, ..., a_r. Sehingga persamaan kurva yang mewakili titik² data dapat diperoleh.