محاسبات عددی

Numerical Analysis

جلسه دهم

برازش منحنی یا Curve Fitting

تفاوت برازش با درون یابی در این است که در درون یابی، چند جمله ای درونیاب بایستی از تمامی نقاط مورد درون یابی عبور کند ولی در برازش الزاماً منحنی برازش شده از تمامی نقاط عبور نمی کند و در واقع بهترین انطباق را بر نقاط پراکندهٔ موجود دارد.

برازش چند جمله ای ها عموماً از نواسانات بالا رنج می برد و گاهاً در مواردی که نقاط به طور یکنواخت صعودی یا نزولی یا ثابت پراکنده شده اند، معادلهٔ چند جمله ای برای نقاط میانی غیر جدولی مقادیر غیر قابل قبولی را ارائه می دهد.

به عنوان مثال چنانچه برای نقاط جدولی زیر که مربوط به تابع $y = \sqrt[3]{x}$ است، چند جمله ای درون یاب را به کمک روش لاگرانژ بدست آوریم، داریم:

x_i	۰	۱	۸	۲۷	۶۴
f_i	۰	۱	۲	۳	۴

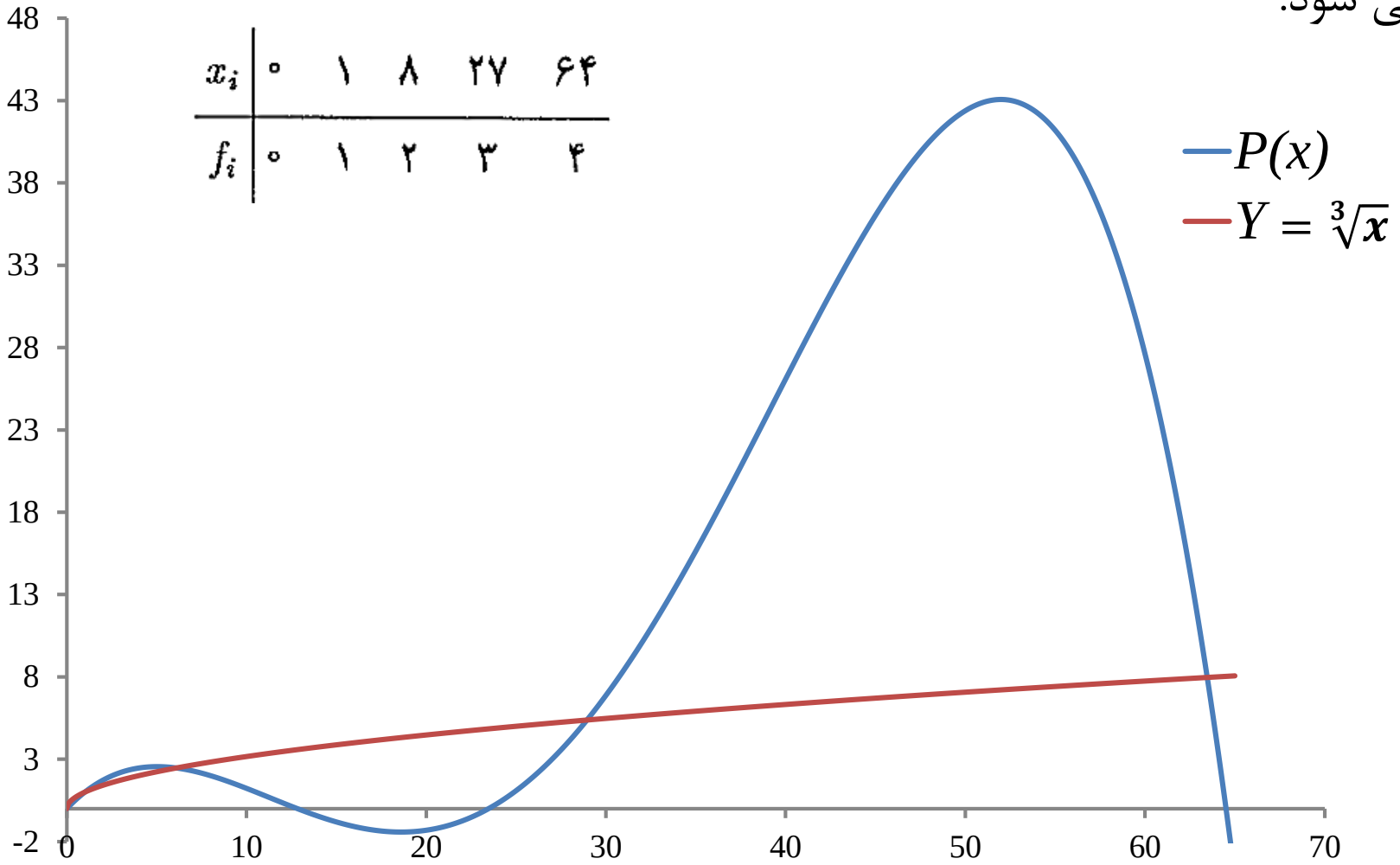
$$P(x) = ۱ \times \frac{x(x-۸)(x-۲۷)(x-۶۴)}{۱(-۷)(-۲۶)(-۶۳)} + ۲ \times \frac{x(x-۱)(x-۲۷)(x-۶۴)}{۸(۷)(-۱۹)(-۵۶)}$$
$$+ ۳ \times \frac{x(x-۱)(x-۸)(x-۶۴)}{۲۷(۲۶)(۱۹)(-۳۷)} + ۴ \times \frac{x(x-۱)(x-۸)(x-۲۷)}{۶۴(۶۳)(۵۶)(۳۷)}$$

چنانچه به کمک این چند جمله ای بخواهیم مقدار تابع $y = \sqrt[3]{x}$ را برای نقطهٔ غیر جدولی $x=20$ بدست آوریم، مقدار $\boldsymbol{P(20) = -1.3139}$ حاصل خواهد شد که این غیر قابل قبول است.

برازش منحنی یا Curve Fitting

با رسم منحنی چند جمله ای درونیاب $P(x)$ بدست آمده از روش لاگرانژ و تابع $y = \sqrt[3]{x}$ در یک دستگاه مختصات در بازهٔ $[0, 65]$ مشاهده می شود که چند جمله ای $P(x)$ فقط در نقاط جدولی جواب درست را بدست می دهد و در سایر نقاط میانی غیر جدولی جوابهای کاملاً پرتی حاصل می شود.

x_i	۰	۱	۸	۲۷	۶۴
f_i	۰	۱	۲	۳	۴



برازش منحنی یا Curve Fitting

بنابراین بهتر است در مواردی که نقاط موجود دارای پراکندگی یکنواخت هستند یا به قولی xها و yها با یکدیگر همبستگی دارند، بهتر است از روش برازش منحنی منحنی بر این نقاط استفاده کنیم.

برای این منظور بایستی:

(الف) با توجه به نوع پراکندگی داده ها، منحنی مناسب برای برازش را انتخاب کنیم

(ب) پارامترهای منحنی برازش را با رویکرد بهترین تقریب محاسبه کنیم

مبنای محاسبات در اندازه گیری پارامترهای منحنی، حداقل سازی مجموع مربعات خطاست (Least Square Error)

معلومات مسئله عموماً به صورت مجموعهٔ نقاط زیر است:

$$(x_i, y_i) \quad i = 1, 2, 3 \ldots n$$

با توجه به نوع پراکندگی نقاط، نوع تابع برازش نیز تعیین می شود که تابعی است از x و تعدادی پارامتر که در اینجا با a_m نشان داده شده است:

$$f(x) = f(x, a_1, a_2, \ldots, a_m)$$

برازش منحنی یا Curve Fitting

با استفاده از حداقل کردن مجموع مربعات خطا در اینجا اقدام به محاسبهٔ پارامترهای منحنی برازش می کنیم:

$$S(a_1, a_2, \ldots, a_m) = \sum_{i=1}^{n} \left[y_i - f(x_i) \right]^2$$

$$r_i = y_i - f(x_i) \equiv residual$$

به منظور حداقل سازی خطا، به ازای هر مولفهٔ مجهول از رابطهٔ بالا مشتق میگیریم:

$$\frac{\partial S}{\partial a_k} = 0, \;\; k = 1, 2, \ldots, m$$

➢برازش خط راست (Regression)

در این حالت نقاط داده به صورت روبرو است: $(x_i, y_i) \quad i = 1, 2, 3 \dots n$

تابع تقریب برازش را نیز به صورت خطی در نظر میگیریم: $f(x) = bx + a$

مشاهده می شود که تابع برازش دارای دو پارامتر a و b است.

بدین ترتیب تابع خطا را بر اساس پارامترهای مجهول تشکیل می دهیم:

$$S(a,b) = \sum_{i=1}^{n} [y_i - a - bx_i]^2$$

به ازای هر مولفهٔ مجهول، تابع خطا را کمینه می کنیم:

$$\frac{\partial S}{\partial a} = \sum_{i=1}^{n} -2(y_i - a - bx_i) = 2\left(-\sum_{i=1}^{n} y_i + na + b\sum_{i=1}^{n} x_i\right) = 0$$

$$\frac{\partial S}{\partial b} = \sum_{i=1}^{n} -2(y_i - a - bx_i)x_i = 2\left(-\sum_{i=1}^{n} x_i y_i + a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2\right) = 0$$

➢برازش خط راست (Regression)

$$\frac{\partial S}{\partial a}=\sum_{i=1}^{n}-2(y_i-a-bx_i)=2\left(-\sum_{i=1}^{n}y_i+na+b\sum_{i=1}^{n}x_i\right)=0$$

$$\frac{\partial S}{\partial b}=\sum_{i=1}^{n}-2(y_i-a-bx_i)x_i=2\left(-\sum_{i=1}^{n}x_iy_i+a\sum_{i=1}^{n}x_i+b\sum_{i=1}^{n}x_i^2\right)=0$$

با تقسیم طرفین معادلات بر 2n و با دانستن $\boxed{\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i}$ و $\boxed{\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i}$ و مرتب نمودن معادلات داریم:

$$2\left(-\sum_{i=1}^{n}y_i+na+b\sum_{i=1}^{n}x_i\right)=0 \xrightarrow{\div\, 2n} a+\bar{x}b=\bar{y}$$

$$2\left(-\sum_{i=1}^{n}x_iy_i+a\sum_{i=1}^{n}x_i+b\sum_{i=1}^{n}x_i^2\right)=0 \xrightarrow{\div\, 2n} a\bar{x}+\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right)b=\frac{1}{n}\sum_{i=1}^{n}x_iy_i$$

برازش منحنی یا Curve Fitting

➢برازش خط راست (Regression)

بدین ترتیب پارامترهای مجهول منحنی برازش عبارتند از:

$$\left.\begin{aligned} a+\bar{x}b&=\bar{y}\\ a\bar{x}+\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right)b&=\frac{1}{n}\sum_{i=1}^{n}x_iy_i \end{aligned}\right\}\begin{cases} \boxed{a=\dfrac{\bar{y}\sum x_i^2-\bar{x}\sum x_iy_i}{\sum x_i^2-n\bar{x}^2}}\\ \boxed{b=\dfrac{\sum x_iy_i-n\overline{xy}}{\sum x_i^2-n\bar{x}^2}} \end{cases}$$

روابط مناسب تر برای کاهش اشتباهات محاسباتی به صورت زیر ارائه می شوند:

$$\boxed{b=\frac{\sum y_i(x_i-\bar{x})}{\sum x_i(x_i-\bar{x})}}\qquad\boxed{a=\bar{y}-\bar{x}b}$$

برازش منحنی یا Curve Fitting

➢برازش خط راست (Regression)

مثال: با توجه به جدول داده های زیر معادلهٔ خط برازش شده بر این داده ها را بدست آورید:

X	0.0	1.0	2.0	2.5	3.0
y	2.9	3.7	4.1	4.4	5.0

$$b=\frac{\sum y_i(x_i-\bar{x})}{\sum x_i(x_i-\bar{x})} \qquad a=\bar{y}-\bar{x}b$$

حل:

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{0.0+1.0+2.0+2.5+3.0}{5}=1.7$$

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i=\frac{2.9+3.7+4.1+4.4+5.0}{5}=4.02$$

$$b=\frac{\sum y_i(x_i-\bar{x})}{\sum x_i(x_i-\bar{x})}=\frac{2.9(-1.7)+3.7(-0.7)+4.1(0.3)+4.4(0.8)+5.0(1.3)}{0.0(-1.7)+1.0(-0.7)+2.0(0.3)+2.5(0.8)+3.0(1.3)}=0.6431$$

$$a=\bar{y}-\bar{x}b=4.02-1.7(0.6431)=2.9267$$

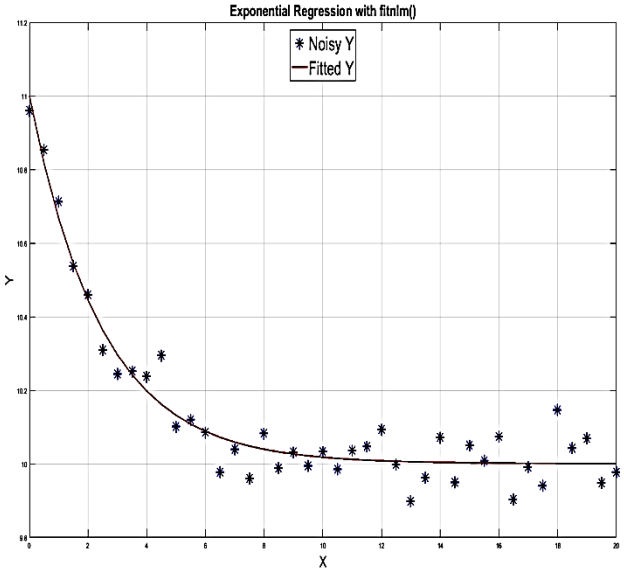
$$y=2.9267+0.6431x$$

➢برازش انواع دیگر منحنی ها

در خیلی از مسائل تقریب چند جمله ای مناسب نیست و ممکن است شکل پراکندگی نقاط داده شده شبیه یک منحنی نمایی، هزلولی یا یک تابع مثلثاتی باشد. بنابراین نیاز است تا انواع دیگری از تقریب های حداقل مربعات مورد استفاده قرار گیرند.

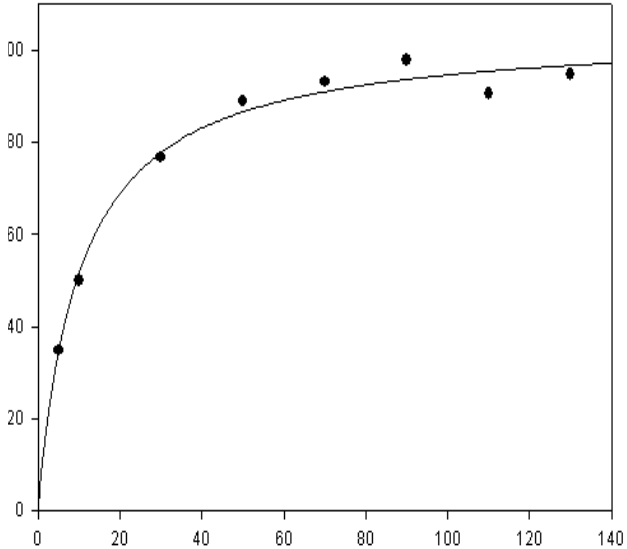
نمایی

Exponential curve fitting



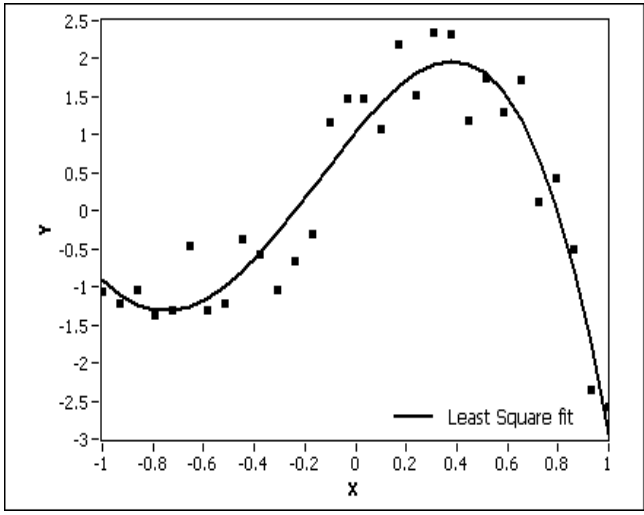
هذلولوی

Hyperbolic curve fitting



چند جمله ای

Polynomial curve fitting



➢برازش منحنی نمایی

هرگاه x_iها و lny_iها رابطهٔ خطی با یکدیگر نشان دادند، بیانگر آن است که x_iها و y_iها با یکدیگر رابطهٔ نمایی دارند و داریم:

$$y = a_0 e^{a_1 x}$$

چنانچه از طرفین ln بگیریم داریم:

$$\ln y = \ln a_0 + a_1 x$$

هرگاه قرار دهیم $A=\ln a_0$ و $z=\ln y$ در این صورت داریم:

$$z = A + a_1 x$$

لذا روش حداقل مربعات برای حداقل کردن مقدار زیر است:

$$S = \sum (\ln y_i - A - a_1 x_i)^2$$

بنابراین معادلات نرمال عبارتند از:

$$\begin{cases} nA + a_1 \sum x_i = \sum \ln y_i \\ A \sum x_i + a_1 \sum x_i^{\ 2} = \sum x_i \ln y_i \end{cases}$$

که از حل آن مقادیر a_1 و A بدست می آیند.

➢برازش منحنی هذلولی

هرگاه با رسم نقاط x_i ها و y_i ها یک هزلولی برای برازش داده ها مناسب تشخیص داده شد داریم:

$$y = \frac{1}{a_0 + a_1 x}$$

که با تغییر متغیر $z=1/y$ داریم:

$$z = a_0 + a_1 x$$

و در آن بایستی عبارت مقابل را حداقل کنیم:

$$S = \sum \left(\frac{1}{y_i} - a_0 - a_1 x_i\right)^2$$

و در این حالت معادلات نرمال عبارتند از:

$$\begin{cases} m a_0 + a_1 \sum x_i = \sum \frac{1}{y_i} \\ a_0 \sum x_i + a_1 \sum x_i^2 = \sum \frac{x_i}{y_i} \end{cases}$$

(که در آن m همان تعداد نقاط است)

از حل معادلات نرمال مقادیر a0 و a1 بدست خواهند آمد که برای نوشتن معادلهٔ منحنی برازش کافی هستند.

➢برازش منحنی چند جمله ای

می خواهیم بر تعداد m نقطه یک چند جمله ای از درجهٔ n برازش کنیم که در آن $n<m$ است.

$$P(x)=\sum_{k=0}^{n} a_k x^k = a_n x^n + \cdots + a_1 x + a_0$$

در اینجا هدف حداقل ساختن عبارت زیر است:

$$S=\sum_{i=1}^{m}[y_i - P(x_i)]^2 = \sum_{i=1}^{m}[y_i - a_n {x_i}^n - \cdots - a_1 x_i - a_0]^2$$

بنابراین مقادیر $a_0, a_1 \ldots a_n$ بایستی طوری محاسبه شوند تا S کمینه گردد.

پس داریم:

$$\frac{\partial S}{\partial a_j}=0, \qquad j=0,1,\cdots,n$$

➢برازش منحنی چند جمله ای

بدین ترتیب معادلات نرمال به صورت زیر بدست می آیند:

$$ma_0 + \left(\sum x_i\right) a_1 + \cdots + \left(\sum x_i^k\right) a_k + \cdots + \left(\sum x_i^n\right) a_n = \sum y_i$$

$$\left(\sum x_i\right) a_0 + \left(\sum x_i^2\right) a_1 + \cdots + \left(\sum x_i^{k+1}\right) a_k + \cdots + \left(\sum x_i^{n+1}\right) a_n = \sum x_i y_i$$

$$\vdots$$

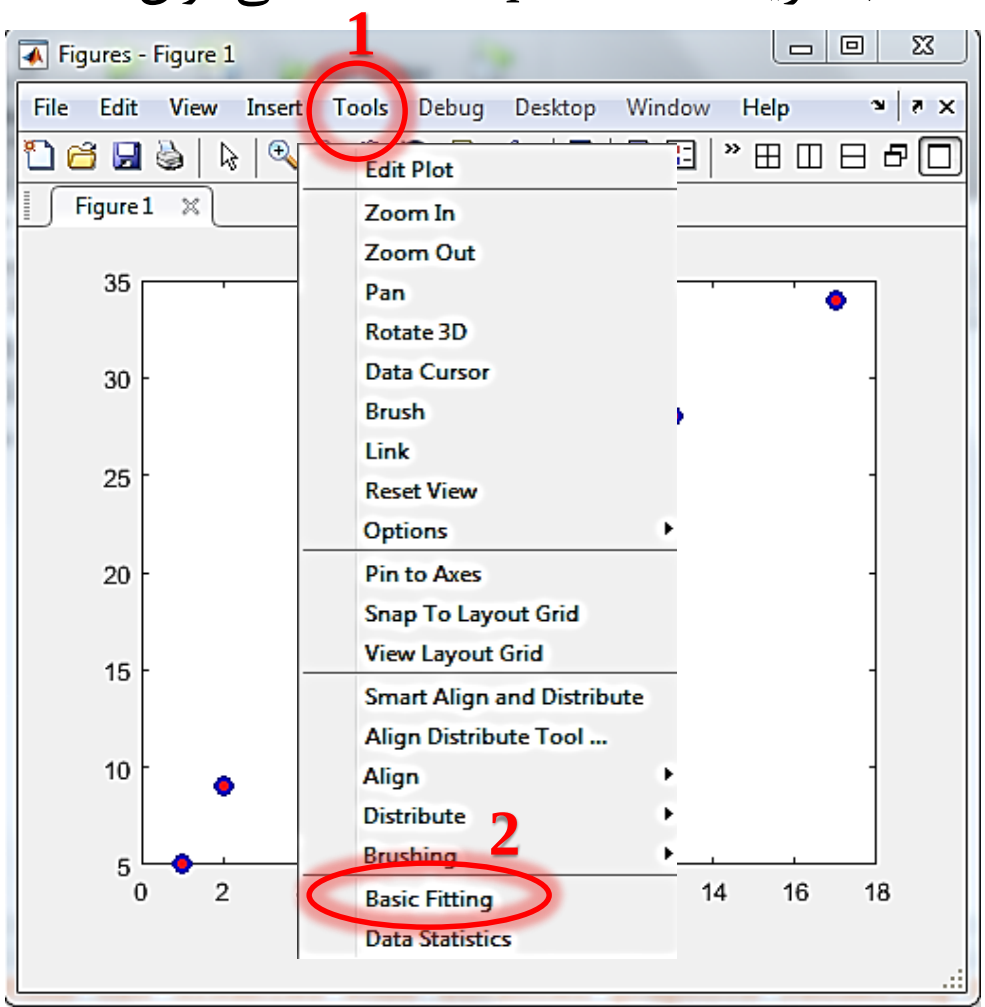
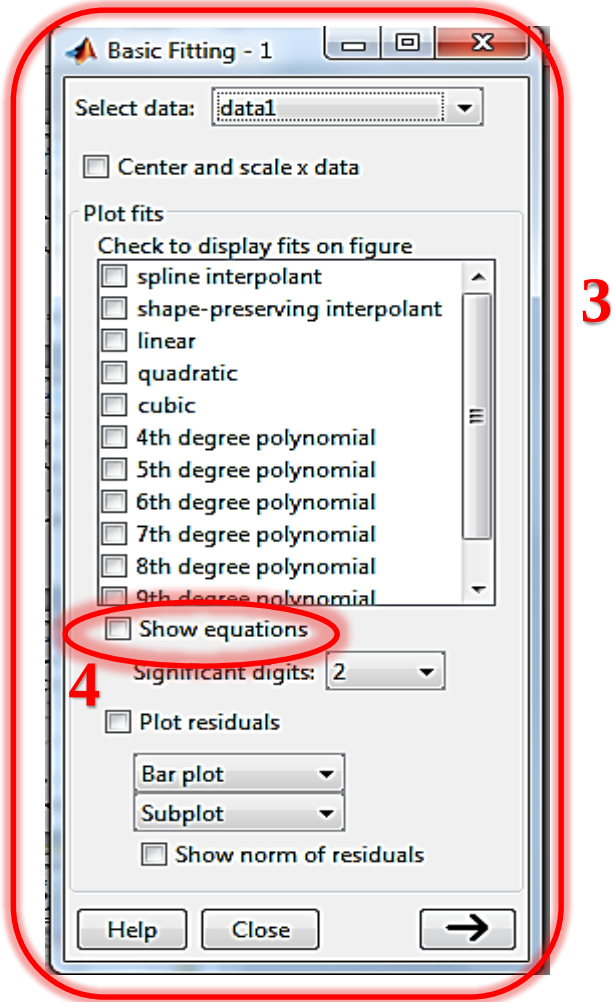
$$\left(\sum x_i^p\right) a_0 + \left(\sum x_i^{p+1}\right) a_1 + \cdots + \left(\sum x_i^{k+p}\right) a_k + \cdots + \left(\sum x_i^{n+p}\right) a_n = \sum x_i^p y_i$$

$$\vdots$$

$$\left(\sum x_i^n\right) a_0 + \left(\sum x_i^{n+1}\right) a_1 + \cdots + \left(\sum x_i^{k+n}\right) a_k + \cdots + \left(\sum x_i^{2n}\right) a_n = \sum x_i^n y_i$$

برای بدست آوردن ضرایب چند جمله ای درجهٔ n میبایست $n+1$ مجهول را از حل $n+1$ معادله بدست آوریم

برازش منحنی یا Curve Fitting

➢برازش به کمک MATLAB

به منظور برازش منحنی در متلب کافی است پراکندگی نقاط داده شده را در یک دستگاه مختصات در کنار یکدیگر رسم کنیم (plot) و سپس از منوی Tools دستور Basic Fitting را انتخاب کرده و در پنجرهٔ باز شده بهترین منحنی که بر داده ها برازش می شود رسم کرده و با انتخاب گزینه show equations می توان معادلهٔ آن را نیز بدست آورد.