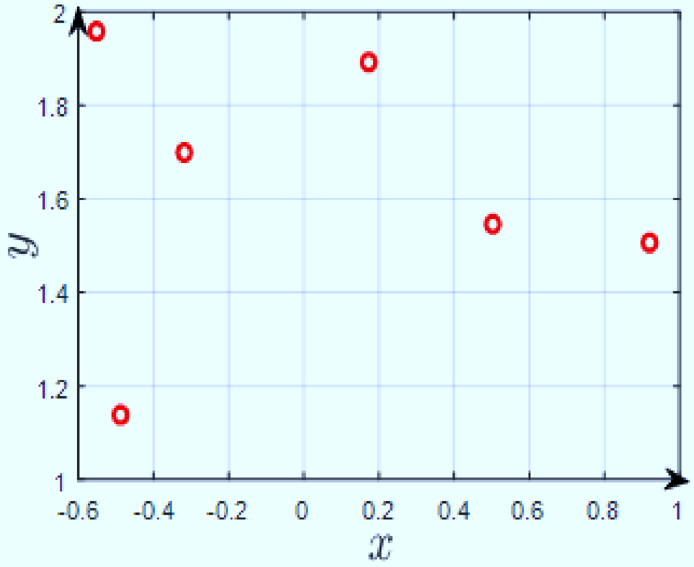
فصل سوم: بخش اول
درونیابی
(Interpolation)

درونیابی یکی از مهمترین مفاهیم در ریاضیات و مهندسی است. در زیر قصد داریم تا یک توصیف هندسی از مساله درونیابی داشته باشیم. فرض کنیم که $(n+1)$ نقطه متمایز $i=0,1...n$ $(x_i, f_i),$ را در صفحه $\mathbb{R}^2$ داریم (مطابق شکل زیر).



نقاط مذکور

یا از تابع معلومی همچون $y = f(x)$ گرفته شده‌اند

یا از تابع خاصی گرفته نشده‌اند و صرفا با برخی اعمال تجربی همچون اندازه‌گیری به دست آمده‌اند

اکنون سوال اینجاست که آیا تابعی همچون $y = G(x)$ را می‌توان یافت که از همه نقاط مذکور عبور نماید؟

یعنی داشته باشیم:

$G(x_i) = y_i,\ \ i = 0,1,\dots,n$

اگر چنین تابعی وجود داشته باشد، آن را **تابع درونیاب** گذرنده از نقاط $(x_i, f_i),\ \ i = 0,1\dots n$ می‌نامیم.

به دلیل خواص بسیار مطلوب چندجمله‌ای‌ها ، در این فصل، تابع درونیاب را یک چندجمله‌ای به صورت:

$$P_n(x) = a_0 + a_1 x + \ldots + a_n x^n$$

فرض می‌کنیم که قرار است از $(n+1)$ نقطه متمایز $i = 0,1 \ldots n$,(x_i, f_i) در صفحه $\mathbb{R}^2$ عبور نماید.
یعنی:

$$P_n(x_i) = f_i, \; i = 0,1,\ldots,n$$

لذا از این پس هر زمان صحبت از درونیابی کردیم، مقصود درونیابی چندجمله‌ای است. (البته درونیاب‌های دیگر همچون درونیاب‌های نمایی، مثلثاتی، تکه‌ای چندجمله‌ای و ... نیز وجود دارند که در این فصل به آن‌ها نمی‌پردازیم)

سوالات اساسی

الف) آیا چنین چندجمله‌ای درونیابی وجود دارد؟ و آیا در صورت وجود یکتا است؟

ب) در صورتیکه این نقاط از تابع معلومی همچون $y = f(x)$ گرفته شده باشند، پیدا کردن این چندجمله‌ای درونیاب اصولا چه فایده‌ای می‌تواند داشته باشد؟

ج) در صورتیکه این نقاط از تابع معلومی گرفته نشده باشند، پیدا کردن این چندجمله‌ای درونیاب چه فایده‌ای می‌تواند داشته باشد؟

برای سوالات دوم و سوم که در واقع سوالات انگیزشی مبحث هستند مصادیقی می‌آوریم و سوال (الف) را نیز در نهایت با یک قضیه پاسخ می‌دهیم.

مصداقی برای سوال (ب): فرض کنیم که حل انتگرال $\int_0^1 \sin(x^2)dx$ مد نظر است. انتگرال فوق به کمک روش‌های تحلیلی قابل حل نیست. اکنون برای حل تقریبی این انتگرال، ایده‌ای را بر اساس مفهوم درونیابی مطرح می‌کنیم. فرض کنیم که یک نمونه گیری از تابع $y = \sin(x^2)$ در بازه $[0,1]$ مطابق با جدول زیر انجام داده‌ایم:

x_i	0	0.2	0.4	0.6	0.8	1
$\sin(x_i^2)$	0	0.039	0.159	0.352	0.597	0.841

لذا مجموعه‌ای از نقاط را در اختیار داریم که از تابعی معلوم گرفته شده‌اند. اکنون چنانچه بتوانیم یک چندجمله‌ای درونیاب طوری پیدا کنیم که از همه نقاط مذکور بگذرد، این بدان معناست که:

$$\sin(x^2) \simeq P_n(x)$$

لذا برای پیدا کردن جوابی تقریبی برای انتگرال فوق می‌توانیم قرار دهیم:

$$\int_0^1 \sin(x^2)dx \simeq \int_0^1 P_n(x)dx$$

و انتگرال یک چندجمله‌ای به سادگی قابل محاسبه خواهد بود.

مصداقی برای سوال (ج): مساله سرشماری را در نظر بگیریم. فرض کنیم جمعیت یک کشور مطابق جدول زیر است:

سال	1360	1370	1380	1390
جمعیت	43709126	51234765	67654907	76523976

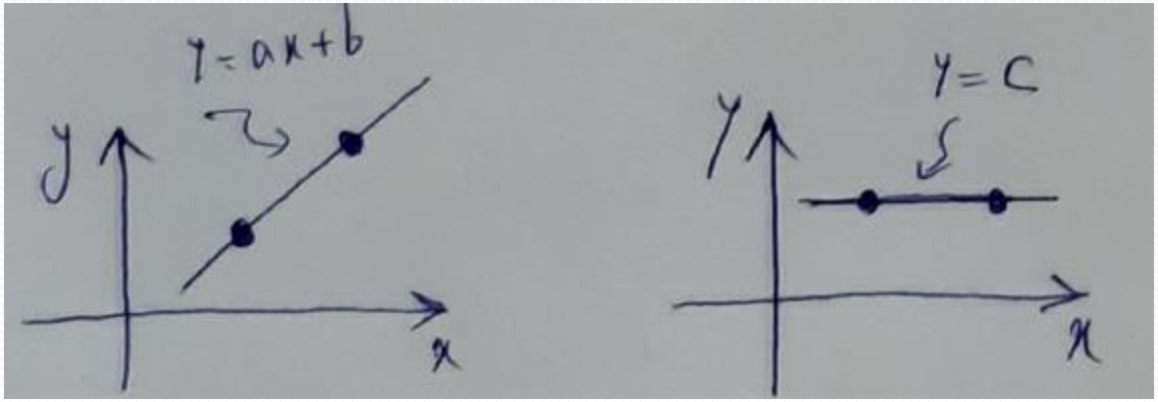
مشخصا اعداد جدول فوق از تابع خاصی گرفته نشده‌اند. اکنون سوال اینجاست که اگر بخواهیم برآوردی از جمعیت مثلا در سال 1387 داشته باشیم، به کمک اعداد جدول فوق این چگونه امکانپذیر است؟

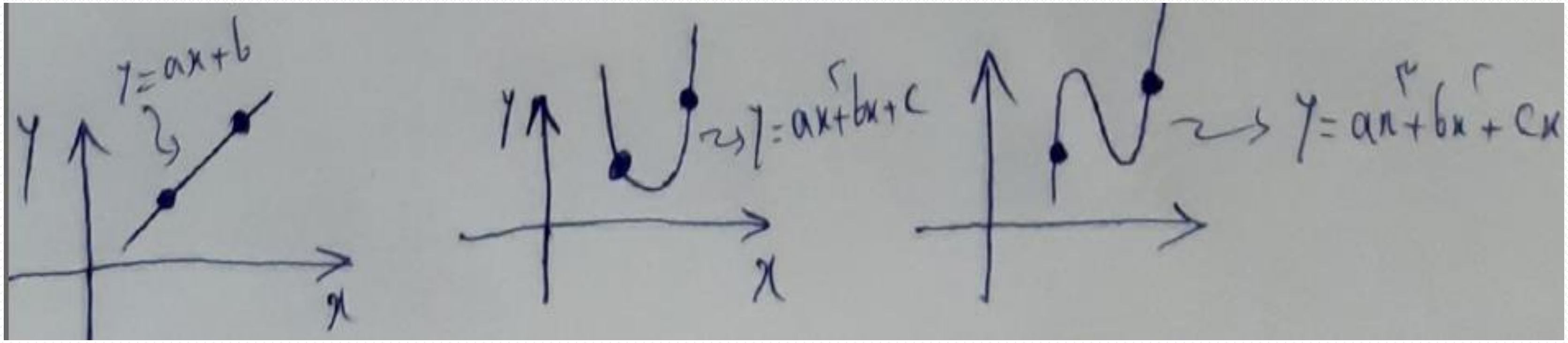
پاسخ می‌تواند به کمک مفهوم درونیابی باشد. چنانچه بتوان یک چندجمله‌ای همچون $y = P_n(x)$ یافت به نحوی که درونیاب جدول فوق باشد، در اینصورت $P_n(1387)$ جوابی تقریبی برای جمعیت در سال 1387 خواهد بود.

پاسخ سوال (الف)، سوال راجع به وجود و یکتایی چندجمله‌ای درونیاب

قضیه (وجود و یکتایی چندجمله‌ای درونیاب): فرض کنیم که $(x_i, f_i),\ i = 0,1,\dots n$ ،$(n+1)$ نقطه متمایز در صفحه $\mathbb{R}^2$ باشند. در اینصورت یک چندجمله‌ای درونیاب حداکثر از درجه n یکتا همچون $y = P_n(x)$ یافت می‌شود به نحوی که:

$$P_n(x_i) = f_i,\ i = 0,1,\dots,n$$





روش‌های پیدا کردن چندجمله‌ای درونیاب

۱ـ روش مستقیم و مشکل ناپایداری محاسباتی: طبیعتا چون قرار است که یک چندجمله‌ای همچون

$$P_n(x) = a_0 + a_1 x + \ldots + a_n x^n$$

داشته باشیم که از نقاط $(x_i, f_i),\ i = 0,1,\ldots,n$ عبور نماید، لذا:

$$\begin{cases} P_n(x_0) = f_0 \Rightarrow a_0 + a_1 x_0 + a_2 x_0^2 + \ldots + a_n x_0^n = f_0 \\ P_n(x_1) = f_1 \Rightarrow a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_n x_1^n = f_1 \\ \vdots \\ P_n(x_n) = f_n \Rightarrow a_0 + a_1 x_n + a_2 x_n^2 + \ldots + a_n x_n^n = f_n \end{cases} \Rightarrow \begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}$$

ماتریس ضرایب فوق به ماتریس واندرموند معروف می‌باشد. متاسفانه دستگاه معادلات خطی $AX = b$ ای که در آن ماتریس A ماتریس واندرموند باشد، مساله‌ای بدحالت محسوب می‌شود. بدحالتی دستگاه فوق به این معناست که خطاهای روندآف در حل این دستگاه چنان تقویت می‌شوند که در جواب نهایی برای ضرایب این دستگاه این خطاها انباشته شده و موجب اصطلاحا ناپایداری محاسباتی می‌گردند!!

۲- روش درونیابی لاگرانژ (Lagrange Interpolation Method): فرض کنیم که $i=0,1,\ldots,n$ ،(x_i,f_i)، همان $(n+1)$ نقطه متمایز در صفحه $\mathbb{R}^2$ می‌باشند. لاگرانژ با تکیه بر چندجمله‌ای‌های زیر:

$$L_0(x)=\frac{(x-x_1)\times(x-x_2)\times\ldots\times(x-x_n)}{(x_0-x_1)\times(x_0-x_2)\times\ldots\times(x_0-x_n)}$$
$$L_1(x)=\frac{(x-x_0)\times(x-x_2)\times\ldots\times(x-x_n)}{(x_1-x_0)\times(x_1-x_2)\times\ldots\times(x_1-x_n)}$$
$$\vdots$$
$$L_n(x)=\frac{(x-x_0)\times(x-x_1)\times\ldots\times(x-x_{n-1})}{(x_n-x_0)\times(x_n-x_1)\times\ldots\times(x_n-x_{n-1})}$$

موسوم به چندجمله‌ای‌های لاگرانژ ، فرمول چندجمله‌ای درونیاب گذرنده از نقاط فوق را به صورت زیر معرفی نمود:

$$P_n(x)=L_0(x)f_0+L_1(x)f_1+\ldots+L_n(x)f_n=\sum_{i=0}^{n}L_i(x)f_i$$

تذکر: می‌توان مشاهده کرد که چندجمله‌ای‌های لاگرانژ دارای خاصیت جالب زیر هستند:

$$L_i(x_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \longrightarrow$$

خاصیت دلتای کرونکر (Kronecker)

و همین خاصیت سبب می‌شود که:

$$P_n(x_0) = L_0(x_0)f_0 + L_1(x_0)f_1 + \dots + L_n(x_0)f_n = 1\times f_0 + 0\times f_1 + \dots + 0\times f_n = f_0$$
$$P_n(x_1) = L_0(x_1)f_0 + L_1(x_1)f_1 + \dots + L_n(x_1)f_n = 0\times f_0 + 1\times f_1 + \dots + 0\times f_n = f_1$$
$$\vdots$$
$$P_n(x_n) = L_0(x_n)f_0 + L_1(x_n)f_1 + \dots + L_n(x_n)f_n = 0\times f_0 + 0\times f_1 + \dots + 1\times f_n = f_n$$

چندجمله‌ای $P(x)$ را که مربوط به تابع جدولی زیر است حساب کنید.

x_i	-۱	۰	۱
f_i	۱	۱	۳

حل: در این مثال $n=۲$ و در نتیجه چندجمله‌ایهای لاگرانژ از درجهٔ دو هستند.
چندجمله‌ایهای لاگرانژ به قرار زیرند:

$$L_۰(x)=\frac{(x-x_۱)(x-x_۲)}{(x_۰-x_۱)(x_۰-x_۲)}=\frac{(x-۰)(x-۱)}{(-۱-۰)(-۱-۱)}=\frac{x^۲-x}{۲}$$

$$L_۱(x)=\frac{(x-x_۰)(x-x_۲)}{(x_۱-x_۰)(x_۱-x_۲)}=\frac{(x+۱)(x-۱)}{(۰+۱)(۰-۱)}=\frac{x^۲-۱}{-۱}$$

$$L_۲(x)=\frac{(x-x_۰)(x-x_۱)}{(x_۲-x_۰)(x_۲-x_۱)}=\frac{(x+۱)(x-۰)}{(۱+۱)(۱-۰)}=\frac{x^۲+x}{۲}$$

از این رو، بنابر فرمول (۲.۴) داریم

$$P(x)=۱\times L_۰(x)+۱\times L_۱(x)+۳\times L_۲(x)=\frac{x^۲-x}{۲}-(x^۲-۱)+\frac{۳(x^۲+x)}{۲}$$

$$P(x)=x^۲+x+۱$$

اشکالات روش درونیابی لاگرانژ:

- محاسبات طولانی و خسته کننده‌ای دارد
- اگر یک یا چند نقطه به نقاط جدولی اضافه گردند، همه محاسبات برای پیدا کردن $L_i(x)$ ها باید از سر گرفته شود
- درجه تابع درونیاب تنها پس از اتمام همه محاسبات مشخص می‌شود و حتی از روی درجه چندجمله‌ای‌های لاگرانژ مشخص نمی‌گردد

مشاهده کردیم که در حالت کلی یک چندجمله‌ای درونیاب روی نقاط $(x_i, f_i),\ \ i = 0,1,\dots,n$ ، به نوعی تابع $y = f(x)$ را تقریب می‌زند. یعنی به ازای یک $x \in [x_0, x_n]$ ، همواره در درونیابی داریم:

$$f(x) \simeq P_n(x)$$

اکنون سوال اینجاست که این تقریب تا چه اندازه دقت دارد؟ پاسخ این سوال را با قضیه زیر می‌دهیم.

<u>قضیه (خطای درونیابی):</u> فرض کنیم $y = P_n(x)$ چندجمله‌ای درونیاب حداکثر از درجه n مربوط به تابع $y = f(x)$ در نقاط متمایز x_0 ، x_1 ، ... ، x_n بوده و تابع f نیز دارای مشتق $(n+1)$ام متناهی باشد. در اینصورت به ازای هر $x \in [x_0, x_n]$ داریم:

$$f(x) = P_n(x) + \frac{(x-x_0)\times(x-x_1)\times\dots\times(x-x_n)}{(n+1)!} f^{(n+1)}(c)$$

که در آن، c عددی در بازه $[x_0, x_n]$ است و البته جای دقیق آن مشخص نمی‌باشد!

تذکر: با توجه به قضیه خطای درونیابی، می‌توان نوشت:

$$E_A(x) = \left|f(x) - P_n(x)\right| = \left|(x - x_0)\times(x - x_1)\times\ldots\times(x - x_n)\right|\times\frac{\left|f^{(n+1)}(c)\right|}{(n+1)!}$$

و از آنجایی که محاسبه $\left|f^{(n+1)}(c)\right|$ عملا امکانپذیر نمی‌باشد (چون جای دقیق c مشخص نیست)، لذا فقط به یک کران برای خطای درونیابی بسنده می‌شود. یعنی چنانچه بتوان:

$$\left|f^{(n+1)}(c)\right| \le M$$

در آن صورت، یک کران برای خطای مطلق درونیابی به ازای نقطه $x \in [x_0, x_n]$ به صورت زیر بدست می‌آید:

$$E_A(x) = \left|f(x) - P_n(x)\right| \le \frac{M}{(n+1)!}\times\left|(x - x_0)\times(x - x_1)\times\ldots\times(x - x_n)\right|$$

تمرین چالشی) آیا طبق کران خطای درونیابی، می‌توان ادعا کرد که با افزایش تعداد نقاط، دقت چندجمله‌ای درونیاب همیشه بالا خواهد رفت؟!!

۳- روش درونیابی نیوتن (Newton Interpolation Method) : نیوتن با استفاده از مجموعه توابع پایه‌ای زیر:

$$\{1,(x-x_0),(x-x_0)(x-x_1),\dots,(x-x_0)(x-x_1)\times\dots\times(x-x_{n-1})\}$$

ادعا کرد که چندجمله‌ای درونیاب گذرنده از $(n+1)$ نقطه $(x_i,f_i),\ i=0,1,\dots,n$ به صورت ترکیبی خطی از اعضای مجموعه فوق و به صورت زیر قابل نوشتن است:

$$P_n(x)=a_0+a_1(x-x_0)+a_2(x-x_0)(x-x_1)+\dots+a_n(x-x_0)(x-x_1)\times\dots\times(x-x_{n-1})$$

که در رابطه فوق ضرایب $\{a_i\}_{i=0}^{n}$ مجهول هستند و باید طوری پیدا شوند که:

$$P_n(x_i)=f_i,\ i=0,1,\dots,n$$

بنابراین باید داشته باشیم:

$$P_n(x_0)=f_0\Rightarrow a_0=f_0$$

$$P_n(x_1)=f_1\Rightarrow a_0+a_1(x_1-x_0)=f_1\Rightarrow a_1=\frac{f_1-f_0}{x_1-x_0}$$

$$P_n(x_2)=f_2\Rightarrow a_0+a_1(x_2-x_0)+a_2(x_2-x_0)(x_2-x_1)=f_2\Rightarrow a_2=\frac{f_2-f_0-\dfrac{f_1-f_0}{x_1-x_0}(x_2-x_0)}{(x_2-x_0)(x_2-x_1)}$$

$$\vdots$$

همانگونه که پیداست، این ضرایب با روالی بازگشتی در حال پیدا شدن هستند. اما هر چه بیشتر جلو برویم، فرمول‌های محاسبه a_3، a_4، ...، a_n پیچیده و پیچیده‌تر می‌شود!

خوشبختانه نیوتن با استفاده از ابزار **تفاضلات تقسیم شده** روال مناسب و سیستماتیکی را برای محاسبه ضرایب $\{a_i\}_{i=0}^{n}$ ارایه نمود. در ادامه این ابزار معرفی خواهد شد.

تعریف: تفاضلات تقسیم شده نیوتنی: فرض کنیم f تابعی باشد که در نقاط $x_0, x_1, \ldots, x_n$ تعریف شده باشد. در این صورت:

- تفاضل تقسیم شده مرتبه صفر تابع f در نقطه دلخواه x_i را با $f[x_i]$ نمایش داده و به صورت زیر تعریف می‌کنیم:

$$f[x_i] = f(x_i) = f_i$$

- تفاضل تقسیم شده مرتبه یک تابع f در نقاط دلخواه x_i و x_j را با $f[x_i, x_j]$ نمایش داده و به صورت زیر تعریف می‌کنیم:

$$f[x_i, x_j] = \frac{f[x_j] - f[x_i]}{x_j - x_i} = \frac{f_j - f_i}{x_j - x_i}$$

- به ازای $k > 0$، تفاضل تقسیم شده مرتبه k بین نقاط x_{i_0}، ...، x_{i_k} را با $f[x_{i_0}, \ldots, x_{i_k}]$ نمایش داده و به صورت زیر تعریف می‌کنیم:

$$f[x_{i_0}, \ldots, x_{i_k}] = \frac{f[x_{i_1}, \ldots, x_{i_k}] - f[x_{i_0}, \ldots, x_{i_{k-1}}]}{x_{i_k} - x_{i_0}}$$

باید توجه داشت که این تفاضلات تقسیم شده نیوتنی صرفا اعدادی هستند که در یک روال بازگشتی تعیین شده و دارای تعبیر هندسی یا فیزیکی خاصی نمی‌باشند.

تذکر: معمولاً یک راهکار جدولی ساده برای تعیین این تفاضلات نیوتن وجود دارد که ما به عنوان نمونه آن را برای ۴ نقطه (x_0, f_0)، (x_1, f_1)، (x_2, f_2) و (x_3, f_3) بصورت زیر توضیح می‌دهیم:

x	f	تفاضلات مرتبه اول	تفاضلات مرتبه دوم	تفاضلات مرتبه سوم
x_0	f_0			
		$f[x_0,x_1]=\dfrac{f_1-f_0}{x_1-x_0}=\square$		
x_1	f_1		$f[x_0,x_1,x_2]=\dfrac{\triangle-\square}{x_2-x_0}=\diamond$	
		$f[x_1,x_2]=\dfrac{f_2-f_1}{x_2-x_1}=\triangle$		$f[x_0,x_1,x_2,x_3]=\dfrac{⏢-\diamond}{x_3-x_0}=⬡$
x_2	f_2		$f[x_1,x_2,x_3]=\dfrac{\bigcirc-\triangle}{x_3-x_1}=⏢$	
		$f[x_2,x_3]=\dfrac{f_3-f_2}{x_3-x_2}=\bigcirc$		
x_3	f_3			

نیوتن در نهایت به کمک این تفاضلات تقسیم شده، فرمول تابع درونیاب را به صورت زیر ارایه داد:

اگر $n+1$ نقطه متمایز $(x_i, f_i),\ i=0,\cdots,n$، داده شده باشند، آن‌گاه چندجمله‌ای درونیاب این نقاط عبارت است از

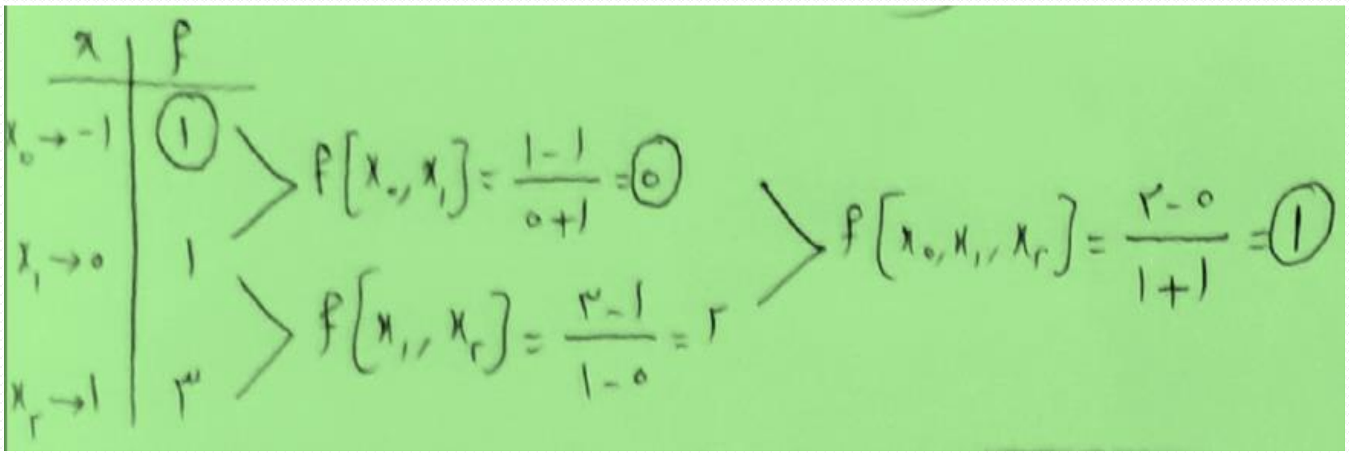
$$p_n(x) = f[x_0]+(x-x_0)f[x_0,x_1]+(x-x_0)(x-x_1)f[x_0,x_1,x_2]+\cdots+$$
$$(x-x_0)\cdots(x-x_{n-1})f[x_0,\cdots,x_n]$$

فرم فشرده چندجمله‌ای درونیاب معرفی شده در قضیه فوق را می‌توان به صورت زیر نوشت

$$p_n(x) = \sum_{i=0}^{n}(x-x_0)\cdots(x-x_{i-1})f[x_0,\cdots,x_i]$$

به سادگی می‌توان تحقیق کرد که فرمول فوق شرایط درونیابی یعنی $P_n(x_i)=f_i,\ i=0,1,\ldots,n$ را برآورده می‌نماید.

مثال) چندجمله‌ای درونیاب گذرنده از نقاط $(-1,1)$، $(0,1)$ و $(1,3)$ را یافته و سپس نقطه $(2,7)$ را به نقاط جدولی اضافه کرده و مجددا تابع درونیاب را بیابید. در مقایسه با روش لاگرانژ چه می‌توان گفت؟



پس چندجمله‌ای درونیاب به کمک جدول تفاضلات تقسیم شده فوق عبارت است از:

$$P_n(x) = 1 + 0 \times (x+1) + 1 \times (x+1)(x-0) = x^2 + x + 1$$

اکنون با اضافه کردن نقطه $(2,7)$ به جدول فوق داریم:

x	f			
-۱	۱			
		۰		
۰	۱		۱	
		۲		۰
۱	۳		۱	
		۴		
۲	۷			

پس چندجمله‌ای درونیاب به کمک جدول تفاضلات تقسیم شده فوق عبارت است از:

$$P_n(x) = 1 + 0\times(x+1) + 1\times(x+1)(x-0) + 0\times(x+1)(x-0)(x-1) = x^2 + x + 1$$

در مقایسه این روش با روش درونیابی لاگرانژ چه می‌توان گفت؟ آیا مشکلات روش لاگرانژ حل می‌شود؟

۴- روش درونیابی بر اساس تفاضلات پیشرو: این روش را باید در واقع حالت خاصی از روش نیوتنی تلقی نمود. در حالتی که نقاط جدولی ما متساوی‌الفاصله، یعنی نقاطی به شکل:

$$x_0,\ x_0+h,\ x_0+2h, \ldots,\ x_0+nh$$

و البته در حالت کلی به صورت، $x=x_0+rh$ باشند، می‌توان نشان داد که فرمول نیوتنی شکل ساده‌تری نیز پیدا خواهد نمود. برای فهمیدن این موضوع در واقع باید در فرمول درونیاب نیوتن از دو چیز تعیین تکلیف شود:

- از تفاضلات تقسیم شده نیوتنی
- از پرانتزهایی که در فرمول نیوتنی به کار رفته‌اند

۱- از تفاضلات تقسیم شده نیوتنی:

$$f[x_0,x_1]=\frac{f_1-f_0}{x_1-x_0}=\boxed{\frac{\Delta f_0}{h}}$$

$$\Delta f_0=\Delta f(x_0)=f(x_0+h)-f(x_0)=f(x_1)-f(x_0)=f_1-f_0$$

$$f[x_0,x_1,x_2]=\frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0}=\frac{\dfrac{f_2-f_1}{x_2-x_1}-\dfrac{f_1-f_0}{x_1-x_0}}{x_2-x_0}=\frac{\dfrac{\Delta f_1}{h}-\dfrac{\Delta f_0}{h}}{2h}=\frac{\Delta f_1-\Delta f_0}{2h^2}$$

$$=\boxed{\frac{\Delta^2 f_0}{2h^2}}$$

$$\Delta^2 f_0=\Delta(\Delta f_0)=\Delta(f_1-f_0)=\Delta f_1-\Delta f_0=(f_2-f_1)-(f_1-f_0)=f_2-2f_1+f_0$$

$$\vdots$$

$$f[x_0,x_1,\dots,x_n]=\cdots=\boxed{\frac{\Delta^n f_0}{n!\,h^n}}$$

۲- از پارامترهایی که در فرمول نیوتنی ظاهر می‌گردند:

$$(x-x_0)=\cancel{x_0}+rh-\cancel{x_0}=\boxed{rh}$$

$$(x-x_0)(x-x_1)=rh\,[\underbrace{x_0+rh-(x_0+h)}_{h(r-1)}]=\boxed{h^2 r(r-1)}$$

$$\vdots$$

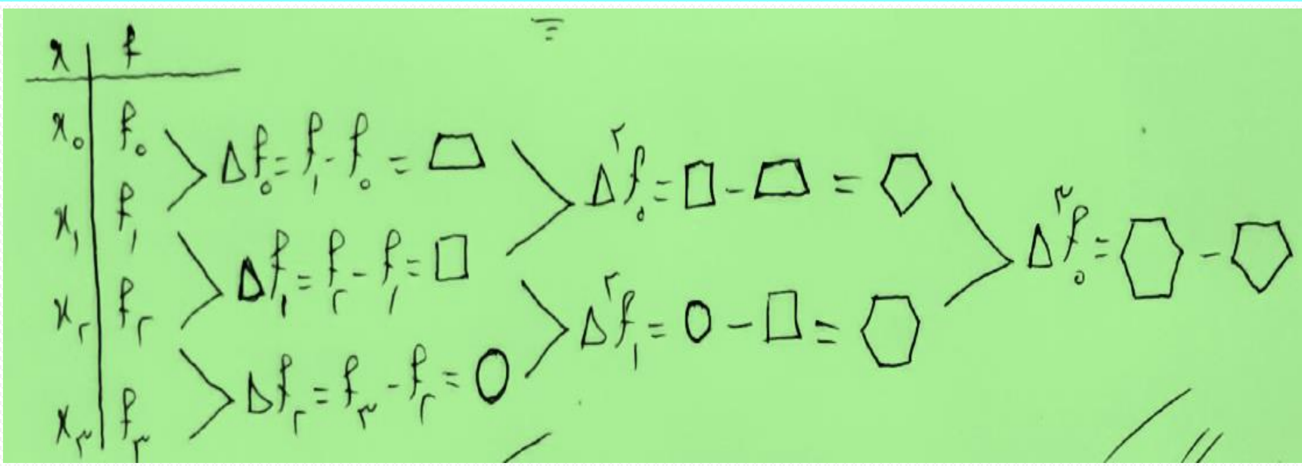
$$(x-x_0)\times\cdots\times(x-x_{n-1})=\cdots=\boxed{h^n r(r-1)(r-2)\times\cdots\times(r-n+1)}$$

با جایگذاری موارد به دست آمده در فرمول نیوتنی در نهایت به یک چندجمله‌ای درونیاب، این بار بر حسب r و به صورت زیر خواهیم رسید:

$$P_n(r) = f_0 + r\Delta f_0 + \frac{r(r-1)}{2!}\Delta^2 f_0 + \dots + \frac{r(r-1)\times\dots\times(r-n+1)}{n!}\Delta^n f_0$$

که البته باید به جای r از $r = \dfrac{x - x_0}{h}$ در آن استفاده کنیم تا بر حسب x بیان شود.

تذکر: همانند روش نیوتن در این جا نیز یک راهکار جدولی سریع برای یافتن تفاضلات پیشرو که در واقع اعدادی بازگشتی هستند، وجود دارد که ما این راهکار جدولی را برای به طور مثال ۴ نقطه جدولی متساوی‌الفاصله توضیح می‌دهیم:



مثال) درونیاب گذرنده از نقاط جدول

x	-1	0	1
f	1	1	3

را به کمک تفاضلات پیشرو بیابید.

x	f		
-1	1		
		$\Delta f_0 = 0$	
0	1		$\Delta^2 f_0 = 2$
		$\Delta f_1 = 2$	
1	3		

$$\Longrightarrow P(r) = f_0 + r\Delta f_0 + \frac{r(r-1)}{2!}\Delta^2 f_0 = 1 + 0\times r + \frac{r(r-1)}{2}\times 2$$

$$= \boxed{1 + r^2 - r} \;\xrightarrow[x=-1+r]{x = x_0 + rh}\; \boxed{r = x+1}$$

پس درونیاب:

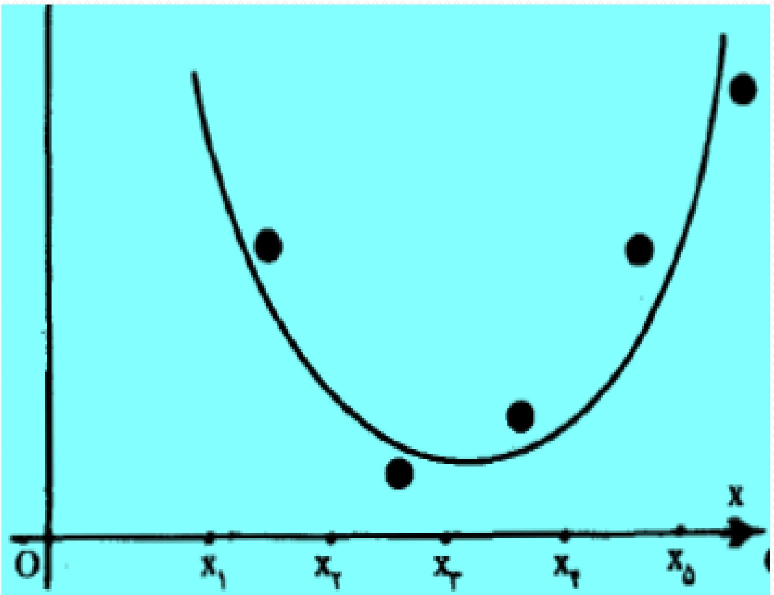
$$P(x) = (x+1)^2 - (x+1) + 1 = \boxed{x^2 + x + 1}$$

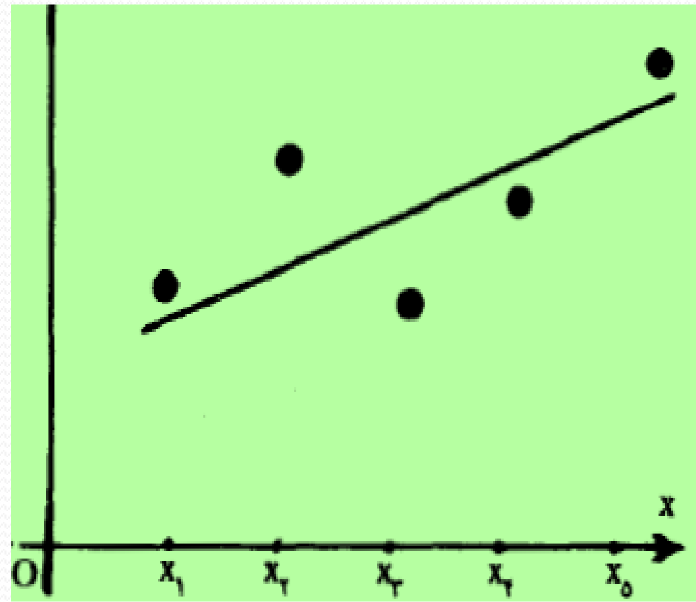
فصل سوم: بخش دوم
برازش منحنی
(Curve Fitting)

کلمه برازش از برازندگی یا برازنده بودن می‌آید. اما مساله برازش منحنی عبارت است از:

یک منحنی پارامتری با ضابطه مشخص و پارامترهای نامشخص همچون $y = f(x; c_1, \dots, c_n)$ ، که در آن $c_1, \dots, c_n$ همان پارامترهای مجهول هستند، داده شده است. (به طور مثال این منحنی پارامتری می‌تواند یک دسته خط به شکل $y = ax + b$ ، یک دسته سهمی به شکل $y = ax^2 + bx + c$ ، یک دسته تابع نمایی به شکل $y = ae^{bx}$ و ... باشد) همچنین مجموعه‌ای از نقاط معلوم در صفحه $\mathbb{R}^2$ مانند $\{(x_1, y_1), \dots, (x_n, y_n)\}$ ، مثلا n نقطه، نیز داده شده است.

اکنون سوال اینجاست که از بین همه منحنی‌های مذکور یعنی $y = f(x; c_1, \dots, c_n)$ ، بهترین منحنی که از نزدیک‌ترین جایگاههای ممکن به این نقاط (و نه الزاما از خود نقاط)، عبور می‌نماید کدام است؟ در واقع هدف یافتن پارامترهای $c_1, \dots, c_n$ در منحنی $y = f(x; c_1, \dots, c_n)$ است، به نحویکه این منحنی از نزدیکترین فواصل ممکن به این نقاط رد بشود.


x
O
x۱
x۲
x۳
x۴
x۵


x
O
x۱
x۲
x۳
x۴
x۵

برای پیدا کردن این بهترین منحنی، نیاز به متر یا معیار جهت سنجش بهترین بودن داریم. طبیعتا معیار ما باید بر اساس فاصله این نقاط تا منحنی (فاصله عمودی) تنظیم گردد. معمولا معیارهای سه گانه زیر در این زمینه موجود است:

$$E_1(c_1,\dots,c_n) = \underset{1 \le i \le n}{\mathrm{Max}} \left| y_i - f(x_i;c_1,\dots,c_n) \right|$$

$$E_2(c_1,\dots,c_n) = \frac{1}{n}\sum_{i=1}^{n} \left| y_i - f(x_i;c_1,\dots,c_n) \right|$$

$$E_3(c_1,\dots,c_n) = \left(\frac{1}{n}\sum_{i=1}^{n} \left(y_i - f(x_i;c_1,\dots,c_n) \right)^2 \right)^{\frac{1}{2}}$$

پارامترهای $c_1,\dots,c_n$ باید طوری تعیین شوند

که هر یک از این سه معیار انتخابی به مینیمم

مقدار خود برسد.

لذا مساله برازش در واقع یک **مساله بهینه‌سازی** (Optimization) به صورت زیر است:

$$\min \ E_i(c_1,\dots,c_n), \quad i=1 \text{ یا } i=2 \text{ یا } i=3$$

از آنجایی که با اطلاعات فعلی ما، ابزار ما در حل مسایل بهینه‌سازی استفاده از مشتق یا همان گرادیان تابع است، لذا با توجه به قدرمطلقی بودن معیارهای E_1 و E_2 ، ما در این درس معیار سوم یعنی E_3 را که تابعی هموار است، برای مینیمم شدن انتخاب می‌کنیم. لذا در این درس مساله برازش عبارت است از:

$$\min \ E_3(c_1,\dots,c_n) = \left(\frac{1}{n} \sum_{i=1}^{n} (y_i - f(x_i;c_1,\dots,c_n))^2 \right)^{\frac{1}{2}}$$

تذکر: باید توجه داشت که برای مینیمم شدن $E_3(c_1,\dots,c_n)$ ، تنها کافی است که حاصلجمع زیر مینیمم گردد:

$$S(c_1,\dots,c_n) = \sum_{i=1}^{n} (y_i - f(x_i;c_1,\dots,c_n))^2$$

لذا مساله برازش منحنی در درس ما به معنای حل مساله بهینه‌سازی زیر است:

$$\min \ S(c_1,\dots,c_n)=\sum_{i=1}^{n}(y_i-f(x_i;c_1,\dots,c_n))^2$$

به همین خاطر، برخی از منابع به جای عبارت برازش منحنی از عبارت <u>**منحنی کمترین مربعات خطا**</u> از این مبحث یاد کرده‌اند.

برای حل مساله فوق، با تشکیل دستگاه معادلات جبری زیر موسوم به <u>**دستگاه معادلات نرمال**</u> و سپس حل آن، مقدار پارامترهای بهینه برای $c_1,\dots,c_n$ به دست خواهد آمد و حل مساله تمام خواهد بود.

$$\begin{cases}\dfrac{\partial S(c_1,\dots,c_n)}{\partial c_1}=0\\ \dfrac{\partial S(c_1,\dots,c_n)}{\partial c_2}=0\\ \vdots\\ \dfrac{\partial S(c_1,\dots,c_n)}{\partial c_n}=0\end{cases}$$

I. <u>**خط کمترین مربعات خطا:**</u> هدف یافتن بهترین خط $y = ax + b$ با معیار سوم است به نحوی که از نزدیکترین جایگاههای ممکن به نقاط $\{(x_i, y_i)\}_{i=1}^{n}$ عبور نماید. لذا در این جا مساله برازش عبارت است از:

$$\min \ S(a,b) = \sum_{i=1}^{n}(y_i - (ax_i + b))^2$$

که برای حل مساله فوق خواهیم داشت:

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0 \\ \dfrac{\partial S}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} \sum_{i=1}^{n} -2x_i(y_i - (ax_i + b)) = 0 \\ \sum_{i=1}^{n} -2(y_i - (ax_i + b)) = 0 \end{cases} \Rightarrow \begin{cases} a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} 1 = \sum_{i=1}^{n} y_i \end{cases}$$

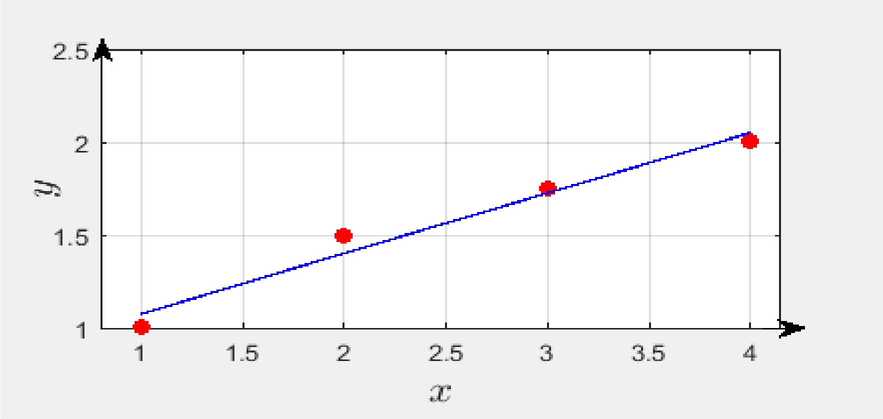
از حل دستگاه معادلات <u>خطی</u> فوق، a و b پیدا شده و حل مساله تمام است.

مثال) خط کمترین مربعات خطا که به داده‌های جدولی زیر، fit می‌شود را بیابید.

x_i	1	2	3	4
y_i	1	1.5	1.75	2

حل: داریم:

$$\begin{cases} \sum_{i=1}^{4} x_i^2 = 1^2 + 2^2 + 3^2 + 4^2 = 30 \\ \sum_{i=1}^{4} x_i = 1+2+3+4 = 10 \\ \sum_{i=1}^{4} 1 = 1+1+1+1 = 4 \\ \sum_{i=1}^{4} x_i y_i = (1\times 1) + (2\times 1.5) + (3\times 1.75) + (4\times 2) = 17.25 \\ \sum_{i=1}^{4} y_i = 1 + 1.5 + 1.75 + 2 = 6.25 \end{cases} \Rightarrow \begin{cases} 30a + 10b = 17.25 \\ 10a + 4b = 6.25 \end{cases} \Rightarrow \begin{cases} a = 0.325 \\ b = 0.75 \end{cases}$$



II. <u>سهمی کمترین مربعات خطا:</u> هدف یافتن بهترین سهمی $y = ax^2 + bx + c$ با معیار سوم است به نحوی که از نزدیکترین جایگاههای ممکن به نقاط $\{(x_i, y_i)\}_{i=1}^{n}$ عبور نماید. لذا در این جا مساله برازش عبارت است از:

$$\min \ S(a,b,c) = \sum_{i=1}^{n}(y_i - (ax_i^2 + bx_i + c))^2$$

که برای حل مساله فوق خواهیم داشت:

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0 \\ \dfrac{\partial S}{\partial b} = 0 \Rightarrow \cdots \Rightarrow \\ \dfrac{\partial S}{\partial c} = 0 \end{cases} \begin{cases} a\sum_{i=1}^{n} x_i^4 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i^2 y_i \\ a\sum_{i=1}^{n} x_i^3 + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} 1 = \sum_{i=1}^{n} y_i \end{cases}$$

از حل دستگاه معادلات <u>خطی</u> فوق، a ، b و c پیدا شده و حل مساله تمام است.

تذکر: چنانچه مشاهده نمودیم، در مواردی که تابع $y = f(x; c_1, \ldots, c_n)$ نسبت به پارامترهای مجهولش خطی باشد، دستگاه معادلات نرمال نهایی حاصل از برازش نیز خطی خواهد بود. واضح است که این اتفاق فقط برای چندجمله‌ای‌ها می‌افتد.

III. **تابع نمایی کمترین مربعات خطا:** هدف یافتن بهترین تابع $y = ae^{bx}$ با معیار سوم است به نحوی که از نزدیکترین جایگاههای ممکن به نقاط $\{(x_i, y_i)\}_{i=1}^{n}$ عبور نماید. لذا در این جا مساله برازش عبارت است از:

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

که برای حل مساله فوق خواهیم داشت:

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0 \\ \dfrac{\partial S}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} \displaystyle\sum_{i=1}^{n} -2e^{bx_i}(y_i - ae^{bx_i}) = 0 \\ \displaystyle\sum_{i=1}^{n} -2ax_i e^{bx_i}(y_i - ae^{bx_i}) = 0 \end{cases}$$

مشاهده می‌شود که این بار با یک دستگاه از معادلات غیرخطی روبرو می‌شویم که دارای حل تحلیلی نیز نمی‌باشد و البته حل عددی آن اگر چه ممکن است، ولی با دردسرهای فراوان ناشی از محاسبه ماتریس ژاکوبی، حدس اولیه مناسب و ... روبرو خواهد بود.

اصولا در مواردی که تابع $y = f(x; c_1, \ldots, c_n)$ نسبت به پارامترهای مجهولش غیر خطی باشد (نظیر همین تابع نمایی)، دستگاه معادلات نرمال نهایی حاصل از برازش نیز غیر خطی خواهد بود.

در چنین مواردی از تکنیکی تحت عنوان **خطی‌سازی** استفاده می‌کنیم.

تعریف: به طور کلی، خطی‌سازی به فرآیندی اطلاق می‌گردد که با پیاده‌سازی آن تابع غیر خطی داده شده در مساله (غیر خطی نسبت به پارامترهای مجهول آن) به یک تابع خطی تبدیل گردد.

مثال) در مساله برازش تابع نمایی $y = ae^{bx}$، با لگاریتم گرفتن از طرفین داریم:

$$\mathrm{Ln}(y) = \mathrm{Ln}(a) + bx \Rightarrow \begin{cases} Y = \mathrm{Ln}(y) \\ X = x \\ c = \mathrm{Ln}(a) \end{cases} \Rightarrow Y = bX + c$$

اکنون برای خط $Y = bX + c$، به سادگی با تشکیل دستگاه معادلات نرمال که دستگاهی خطی و به صورت:

$$\begin{cases} b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} 1 = \sum_{i=1}^{n} y_i \end{cases} \Rightarrow \begin{cases} b = ? \\ c = ? \end{cases}$$

می‌باشد، b و c را یافته و در نهایت از $c = \mathrm{Ln}(a)$ می‌توان نتیجه گرفت که $a = e^c$ و حل مساله برازش بدون مواجهه با دستگاه غیر خطی تمام خواهد بود.

مثال) تابع نمایی $y = ae^{bx}$ کمترین مربعات خطا را که داده‌های جدول زیر را برازش می‌کند، بیابید:

x_i	1	3	4	6	9	15
y_i	4	3.5	2.9	2.5	2.75	2

حل: ابتدا به کمک $\begin{cases} Y = \mathrm{Ln}(y) \\ X = x \end{cases}$ داریم:

$X_i = x_i$	1	3	4	6	9	15
$Y_i = \mathrm{Ln}(y_i)$	1.39	1.25	1.06	0.92	1.01	0.69

اکنون:

$$\begin{cases} \sum_{i=1}^{6} X_i^2 = 368 \\ \sum_{i=1}^{6} X_i = 38 \\ \sum_{i=1}^{6} X_i Y_i = 34.34 \Rightarrow \begin{cases} 368b + 38c = 34.34 \\ 38b + 6c = 6.32 \end{cases} \Rightarrow \begin{cases} b = -0.045 \\ c = 1.34 \Rightarrow a = e^c = 3.82 \end{cases} \\ \sum_{i=1}^{6} 1 = 6 \\ \sum_{i=1}^{6} Y_i = 6.32 \end{cases}$$

بنابراین در نهایت بهترین تابع نمایی کمترین مربعات خطا در این مثال، تابع $y = 3.82e^{-0.045x}$ خواهد بود.