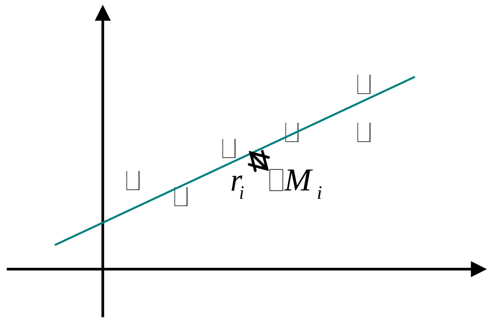
תורת הקירובים

קו רגרסיה:

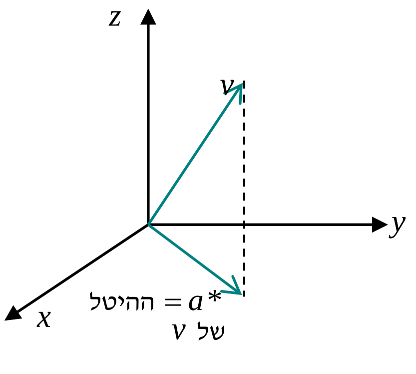
$M_0(x_0,y_0)$
$M_1(x_1,y_1)$
$\vdots$
$M_n(x_n,y_n)$



$$r_i=|y_i-ax_i-b|$$
$$\sum r_i^2\to\min$$
$$\left\|f(x)-P_n^*(x)\right\|_{C[a,b]}\to\min$$
$$\left\|f(x)-L_n(x)\right\|_{C[a,b]}\le(1+\mu)dist\left(f,P_n,C[a,b]\right)$$

נדבר על השיטה הטובה ביותר לקירובים, הן מבחינת יעילות והן מבחינת מציאת מינימום.

שיטה ראשונה: ריבועים מזעריים, *Least-square Method*



יש הרבה וקטורים שונים a שאפשר ליצור מהם
ומווקטור v מישור. a^* הוא המקיים $\left\|x-a^*\right\|=\min$
במקרה שלנו, אותו a^* הוא ההיטל.

מרחב נורמי μ הוא מרחב ללא מכפלה סקלרית. $\forall x\in\mu:\exists\|x\|$. נניח ש-
$A=span\{e_1,e_2,\dots,e_n\}$ מרחב וקטורי. אז $a^*\in A$ יהיה הקירוב הטוב ביותר עבור $x\subset\mu$ אם
$\exists a^*\in A:\|x-a^*\|\to\min$

בדוגמא למעלה: $x=\{v\}$, $A=span\{(1,0,0),(0,1,0)\}$. צריך לבנות את a^*.

$(\mu,\langle x,y\rangle)$ מרחב עם מכפלה סקלרית. $\langle x,y\rangle=$ מ"פ (ליניארית בשני המשתנים, $(x,x)=\|x\|^2$)

מכפלה סקלרית: $(v_1,v_2,v_3)\times(w_1,w_2,w_3)=v_1w_1+v_2w_2+v_3w_3$

אם שני וקטורים מאונכים, אז המכפלה הסקלרית שלהם היא 0! ולכן, הקירוב הטוב ביותר יהיה ההיטל של וקטור v.
מרחב נורמי נוסף:

$$\mu \times \mu \to \mathbb{R}$$

$$\mu = L_2(a,b)$$

$$\langle f(x), g(x) \rangle = \int_a^b f(x)\cdot g(x)dx$$

$$\|f\|_{C[a,b]} = \max_{a\le x\le b}|f(x)|$$

$$\|f\|_{Lp[a,b]}^p = \int_a^b f^p(x)\,dx \longleftarrow p=2$$ בפרט עבור

עבור המרחב $\mathbb{R}^n$ נציג מספר אופציות לנורמות:

$$\|x\|_1 = \sum_{i=1}^n |x_i|$$

$$\|x\|_2^2 = \sum_{i=1}^n |x_i|^2 \Rightarrow \langle x,y \rangle = \sum_{i=1}^n x_i y_i$$

$$\|x\|_p^p = \sum_{i=1}^n |x_i|^p$$

$$\|x\|_\infty = \max_{1\le i\le n}|x_i|$$

הערה: שני וקטורים מאונכים זה לזה $\langle v,w \rangle = 0 \Leftrightarrow$

שתי פונקציות מאונכות זו לזו בקטע $[a,b]$ $\Leftrightarrow \int_a^b f(x)g(x)dx = 0$

רגרסיה:

צריך: $S = \sum_{i=1}^n |y_i - ax_i - b|^2 \to \min$

נפתח:

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

$$0 = \frac{\partial S}{\partial a} = 2\sum_{i=1}^n (y_i - ax_i - b)$$

נציג זאת במטריצה:

$$\begin{pmatrix} -\sum x_i^2 & -\sum x_i \\ -\sum x_i & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i \end{pmatrix}$$

כאשר מנסים לפתור $Ax = y$, $v_1x_1 + v_2x_2 + \ldots + v_nx_n = y \Leftrightarrow Ax = y$
אם אין כזה וקטור x במרחב שלנו, נחפש את הווקטור הקרוב ביותר לפתרון, ע״י $\|Ax - y\| \to \min$. אנו בעצם מנסים למצוא קירוב לפונקציה $a^*\cdot x_i + b^* = y_i$, או:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\cdot a^* + \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}\cdot b^* = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

משפט: כדי למצוא פתרון טוב ביותר ל- $Ax = y$ מספיק למצוא את הפתרון ל- $A^tA\cdot x^* = A^t\cdot y$

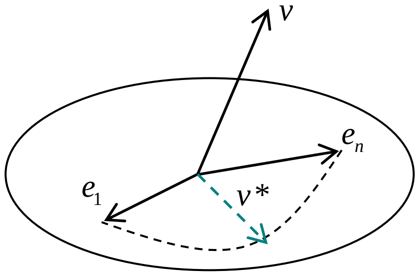
$$\begin{pmatrix} x_1 & \dots & x_n \\ 1 & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = A^t \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \Rightarrow \begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$$

עבור פרבולה:

$$y_i - a x_i^3 - b x_i - c = 0$$

$$\begin{pmatrix} x_1^2 & x_1 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

אולם, כיצד נמצא את ההיטל של וקטור v על תת-מרחב A?
הקירוב הטוב ביותר להיטל יהיה
$v^* = \sum_{i=1}^{n} (v \cdot e_i) \cdot \vec{e}_i$, כאשר $(e_i, e_j) = \delta_{ij}$, בסיס
אורתוגונאלי. אם הבסיס אינו אורתוגונאלי
נשתמש בגרהם-שמידט.



ריבועים זעירים לא-ליניאריים. (Non-linear least square fitting)

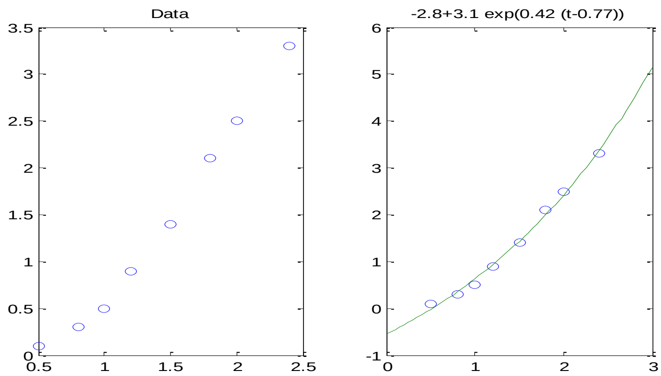
Example: *Fit the model function* $f(t) = a + b\, e^{\omega(t-t_0)}$ *to experimental data:*

t	0.5	0.8	1.0	1.2	1.5	1.8	2.0	2.4
y	0.1	0.3	0.5	0.9	1.4	2.1	2.5	3.3

This yields an ***over determined non-linear system*** *(4 unknowns –* a, b, ω, t_0 *, and 8 equations) which can be solved by demanding that residuals between the measurements* y_i *and the model, namely* $f(t_i)$ *be small as possible, meaning that:* $\| f(t_i) - y_i \|_2 \to \min$.

MATLAB PROGRAM:

```
t=[0.5 0.8 1 1.2 1.5 1.8 2 2.4]';
y=[0.1 0.3 0.5 0.9 1.4 2.1 2.5 3.3]';
subplot(1,2,1), h=plot(t,y,'o'), title('Data')
a=0.7; b=1.7; w=2.; t0=0.2;
x=[a b w t0]';                                  % initial guess
n=size(t,1); iter=0;  xnorm=1.;
while xnorm>1E-6 & iter<10
 u=w*(t-t0); f=a+b*exp(u)-y;
 J=[ones(n,1) exp(u) b*(t-t0).*exp(u) -w*b*exp(u)]; % Jacobian
 dx=-J\f; x=x+dx;                               % Gauss-Newton
 xnorm=norm(dx); iter=iter+1;
 a=x(1); b=x(2); w=x(3); t0=x(4);
end
tt=(0:0.05:3)'; Ft=a+b*exp(w*(tt-t0));
subplot(1,2,2), plot(t,y,'o',tt,Ft);
str=sprintf('%0.2g+%0.2g exp(%0.2g (t-%0.2g))',a,b,w,t0);
title(str)
```



פולינומים אורתוגונאליים

דוגמא: ניקח בסיס $1,x,x^2,...,x^n$ במרחב עם מחפלה סקלרי $L_2[-1,1]$. אז $f(x),g(x)$

פולינומים אורתוגונאליים אם $\int_{-1}^{1} f(x)g(x)dx=0$. נבדוק מתי $e_1=1,e_2=x,e_3=x^2$

אורתוגונאליים לפי לג׳ינדר $(Legendre)$

$$\int_{-1}^{1} 1\cdot xdx=0$$

ולמרות זאת, $\int_{-1}^{1} 1\cdot x^2dx\neq 0$. אנו צריכים לבחור וקטור אחר שיתאים לבסיס האורתוגונאלי:

$x^2+\alpha x+\beta$. את α,β נמצא ע״י שתי משוואות:

$$\int_{-1}^{1}\left(x^2+\alpha x+\beta\right)\cdot 1dx=0$$

$$\int_{-1}^{1} x^3+\alpha x^2+\beta xdx=0$$

באופן כללי, נוכל למצוא וקטורים לבסיס אורתוגונאלי ע״י $\mu(x)$, פולינום נוסף, שיאזן את

האינטגרל $\int_a^b f(x)g(x)\mu(x)dx$ ל-0. $\mu(x)$ נקראת משקולת. ואז:

$L_{2,\mu(x)}=\int_{-1}^{1} f(x)g(x)\mu(x)dx$. למשל: $\int_{-1}^{1}\frac{f(x)g(x)}{\sqrt{1-x^2}}dx$, או כמו בפולינום צ׳ביצ׳ב:

$$T_1(x)=1$$
$$T_2(x)=x$$
$$\vdots$$
$$T_n(x)=\dots$$
$$\int_{-1}^{1}\frac{T_i(x)T_j(x)}{\sqrt{1-x^2}}dx=0,\quad \forall i\neq j$$

דוגמא: $f(x+2\pi)=f(x)$ $[-\pi,\pi]$. נבנה בסיס:

$A = span\left\{ \frac{1}{2\pi}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos kx}{\sqrt{\pi}}, \frac{\sin kx}{\sqrt{\pi}}, \ldots \right\}$

$\int_{-\pi}^{\pi} \cos k\pi \cdot \cos m\pi dx = 0$

$\int_{-\pi}^{\pi} \cos^2 k\pi dx = \pi \longleftarrow$ לא נרמלנו

הקירוב הכי טוב במרחב הפונקציות המחזוריות יהיה:

$f(x) = \sum_{i=1}^{n} \int_{-\pi}^{\pi} f(x) e_i(x) dx \cdot e_i(x)$

$f(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\pi}^{\pi} f(x) dx + \frac{1}{\sqrt{\pi}} \cdot \left[\int_{-\pi}^{\pi} f(x) \cdot \cos x dx \right] \cdot \cos x$

$\Rightarrow f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x \longleftarrow$ מקדמי פורייה

הערה: $C[a,b]$ מרחב נורמי אבל לא מרחב עם מחפלה סקלרי. $L_2[a,b]$ נורמי שאפשר להתאים לו מכפלה סקלרית.

אם e_i אורתוגונאלי אז הנוסחה היא $\delta^* = \sum_{i=1}^{n} \frac{(v, e_i) \cdot \vec{e}_i}{(e_i, e_i)}$

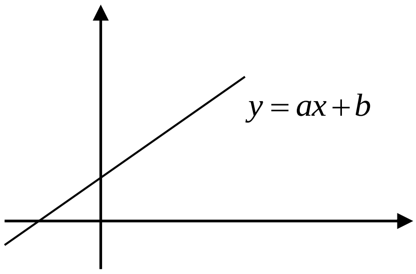
לסיכום:

נחפש את הקירוב הטוב ביותר למציאת v^* כך ש- $\|v - v^*\| \to \min$, כאשר v^* במרחב $(\mu, \langle x, y \rangle)$ עם מכפלה סקלרית (ולא נורמית!). ולכן, לא נמצא בוודאות v^* כך ש- $\|v - v^*\| = 0$.

ניקח תת-מרחב $A = span\{e_1, \ldots, e_n\}$ ונמצא $a^* \in A$ שעבורו $\|v - a^*\| \to \min$ ע"י רישום $e_1, \ldots, e_n$ כווקטור אורתונורמאלי ופתרון ע"י $a^* = \sum_{i=1}^{n} \frac{(v, e_i) \cdot \vec{e}_i}{(e_i, e_i)}$

<u>קו רגרסיה:</u>

x	y
-1	1
0	0
1	1



צריך ש- $S_2(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$

<u>דרך 1</u>: לגזור, להשוות ל-0 ולמצוא את a, b

<u>דרך 2</u>: נגדיר 3 וקטורים:

$\vec{1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad X = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \quad Y = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}$

$\forall 0 \le i \le n : ax_i + b = y_i$

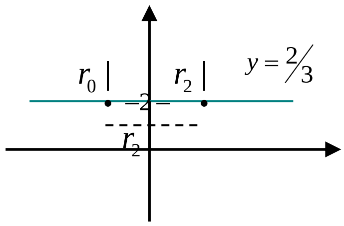
$aX + b \cdot \vec{1} = Y$

נקבל שאין באפשרותנו למצוא במדויק את וקטור Y, ולכן נבצע:

$$Y^*=\frac{(Y,X)}{(X,X)}\cdot X+\frac{(Y,\vec{1})}{(\vec{1},\vec{1})}\cdot\vec{1}=\frac{2}{3}\cdot\vec{1}$$

בדוגמא הנ"ל

ולכן $a^*=0, b^*=\frac{2}{3}$. $\vec{1}, X$ אורתוגונאליים ולכן $(\vec{1},X)=x_0+x_1+x_2=0$



$$r_0=r_2=\frac{1}{3},\ r_1=\frac{2}{3}$$

$$S_2=\frac{1}{9}+\frac{1}{4}+\frac{1}{9}=2/9$$

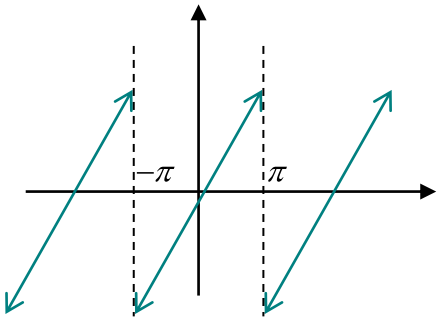
<u>דרך 3:</u> $S_1(a,b)=\sum_{i=0}^{n}|y_i-ax_i-b|\to\min$. פה אין מכפלה סקלרית.

טורי פורייה:

$$f(x)\sim T_n(x)$$

$$f(x+2\pi)=f(x)$$

$$x_\pi=\begin{cases}x & -\pi<x<\pi\\ x-2\pi & \pi<x<3\pi\end{cases}$$



$$T_n(x)=\frac{a_0}{2}+a_1\cos x+b_1\sin x+\ldots+a_n\cos(nx)+b_n\sin(nx)$$

(כי הם אורתוגונאליים) $\int_{-\pi}^{\pi}\cos(kx)\cdot\sin(mx)\,dx=0$

$$\int_{-\pi}^{\pi}\cos(kx)\cdot\cos(mx)\,dx=0\qquad k\neq m$$

$$f(x)\sim T_n(x)=\frac{a_o}{2}+\sum_{i=1}^{n}\left(a_n\cos(nx)+b_n\sin(nx)\right)\longleftarrow$$ קירוב טריגונומטרי

$$a_n=\frac{1}{\pi}\cdot\int_{-\pi}^{\pi}f(x)\cos(nx)\,dx=\frac{1}{\pi}\int_{-\pi}^{\pi}x\cdot\cos(nx)=0$$

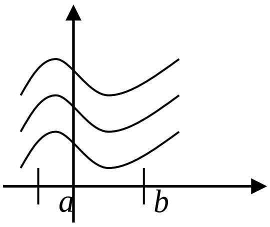
$$b_n=\frac{1}{\pi}\cdot\int_{-\pi}^{\pi}f(x)\sin(nx)\,dx=\frac{1}{\pi}\int_{-\pi}^{\pi}x\cdot\sin(nx)=\left(\frac{-1}{n\pi}x\cos(nx)\right)\Bigg|_{-\pi}^{\pi}+\frac{1}{n\pi}\int_{-\pi}^{\pi}\cos nx\,dx$$

$$=\frac{-2}{n}\cos(n\pi)=-\frac{2}{n}\cdot(-1)^n$$

$$\Rightarrow f(x)=\sum_{i=1}^{n}\frac{(-1)^{n+1}}{n}\cdot\sin(nx),\quad E_2=\left\|f(x)-T_n(x)\right\|_{L_2(-\pi,\pi)}$$

מתי כדאי להגדיר אינטגרל עם משקל? $\leftarrow$ $(\mu(x)\geq 0)\,\|f\|^2_{L_2(a,b,\mu(x))}=\int_a^b f^2(x)\mu(x)\,dx$

למשל עבור פולינום צ'ביצ'ב: ניקח $\int_{-1}^{1}\frac{f^2(x)}{\sqrt{1-x^2}}dx \to \min$. באופן כזה נוכל לשים דגש על הקטע החשוב ביותר. במקרה ספציפי זה, כדאי לתת יותר חשיבות לנקודות בקצות הקטע.



ההבדל בין נורמה L_2 לבין C

$y=f(x), \left\| f(x)-P_n(x)\right\|_{C[a,b]} \le \varepsilon$ תמיד נשאף עבור

$f(x)-\varepsilon \le P_n(x) \le f(x)+\varepsilon$ $\quad P_n(x)\to$

$\left\| f(x)-g(x)\right\|_{L_2[a,b]} = \int_a^b \left(f(x)-g(x)\right)^2 dx \underset{y=f(x)}{=} s_1+s_2+s_3$ כשיש הרבה תנודות, נעדיף את L_2. נורמה של C היא יותר טובה,אבל יותר קשה לחשב אותה.

$$\|f\|_{L_2(a,b)} = \int_a^b \left|f(x)\right|^2 dx \le \left(\max\left\{f^2(x)\right\}\right)^2 \cdot \int_a^b dx = (b-a)\cdot\|f\|_C^2$$

$\Leftarrow$ הדיוק הכי טוב ב-C הוא גם הדיוק הכי טוב ב-L_2.

<u>אינטגרציה נומרית</u>

כיצד מקרבים $\int_a^b f(x)dx$?

אופרטור ליניארי מוצג לרוב כך: $L(xf+\mu g)=\lambda Lf+\mu Lg$. אינטגרל הוא גם פונקציונל ליניארי.

כיצד שומרים אופרטור ליניארי? $Lf(x) \sim \sum_{i=1}^{n} w_i f(x_i)$

משקולות

צמתים

<u>לדוגמא</u>: חישוב נומרי של נגזרת:

$$\frac{\partial f(x)}{\partial x} \sim f_h(x) := \frac{f(x+h)-f(x)}{h}$$

$$x_1 = x+h,\ w_1=\frac{1}{h},\ x_2=x,\ w_2=-\frac{1}{h},$$

ואז הנוסחה של האובייקט תהיה:

$$I(f)=\int_a^b f(x)\mu(x)dx=\sum_{i=1}^{n} w_i f(x_i)=I^h(f)$$

$I^h(f)$ היא מדרגת דיוק אלגברי m אם:

$$I_1 = I_h(1)$$
$$I_x = I_x(1)$$
$$\vdots$$
$$I\left(x^m\right)=I_h\left(x^m\right)$$
$$I\left(x^{m+1}\right)\ne I_h\left(x^{m+1}\right)$$

(כלומר במעלה $m+1$ כבר שתי נוסחאות לא זהות).

ישנן שיטות קירוב שונות, כגון $\frac{1}{2h}\left[f(x+h)-f(x-h)\right]$. נבדוק את רמת הדיוק של שני האלגוריתמים שראינו:

$f(x)$	$f'(x)$	$\frac{1}{h}\left[f(x+h)-f(x)\right]$	$\frac{1}{2h}\left[f(x+h)-f(x-h)\right]$
1	0	0	0
x	1	$\frac{1}{h}\left[(x+h)-x\right]=1$	1
x^2	$2x$	$\frac{1}{h}\left((x+h)^2-x^2\right)=2x+h$	$\frac{1}{2h}\left[(x+h)^2-(x-h)^2\right]=2x$

אין דיוק אלגברי
של 2

ומכאן שתמיד עדיף את הקירוב $\frac{1}{2h}\left[f(x+h)-f(x-h)\right]$ עבור דרגת קירוב גבוהה.

שיטת תרבעה של גאוס

נרצה למצוא x_i, w_i כך שהדיוק האלגברי יהיה הכי טוב.

$$\int_a^b P_n(x)\cdot P_m(x)\mu(x)dx=0 \qquad (m\neq n)$$

$$P_0=1, P_1(x),\ldots$$

$\mu(x)$	a,b	$P_n(x)$
1	$-1,1$	$Legendre$

ניקח בתור צמתים שורשים של פולינום אורתוגונאלי. $P_n(x_i)=0$ בנבנה נוסחה

$$\int_{-1}^{1}\frac{f(x)}{\sqrt{1-x^2}}dx\sim w_1 f(x_1)+w_2 f(x_2)=w_1 f\left(\frac{-1}{\sqrt{2}}\right)+w_2 f\left(\frac{1}{\sqrt{2}}\right)$$

($x_{1,2}=\pm\frac{1}{\sqrt{2}}$ נבעו מהשורשים של $T_2(x)=2x^2-1$, לפי פולינום צ'ביצ'ב עבור $w_{1,2}=\frac{\pi}{2}$).

במקום $f(x)$ אפשר לבנות פולינום של לגראנז' $f(x)\cong L_n(x)$, כאשר $x_1,\ldots,x_n$ הם שורשים של הפולינום האורתוגונאלי $P_n(x_i)=0$.

$$f(x)\sim L_n(x)=\sum_{i=1}^{n}w_i(x)y_i \qquad \left(w_i(x)=\prod_{i\neq j}\frac{(x-x_j)}{(x_i-x_j)}\right)$$

$$W_i=\int_a^b w_i(x)\mu(x)dx$$

$\mu(x)$ הן משקולות לתרבוע גאוס שנועדו להגדיל את הדיוק ולהביאו למקסימום האפשרי.

נניח שבנינו פולינומים אורתוגונאליים:

$$\int_a^b P_i(x)P_j(x)\mu(x)dx=0 \quad i\neq j$$

$$\int_a^b P_n(x)\mu(x)dx=\int_a^b a_0\mu(x)dx+\sum_{i=1}^{n}a_i\int_a^b P_i(x)\mu(x)dx=a_0\int_a^b\mu(x)dx=a_0\cdot c$$

$$P_n(x)=a_0+a_1p_1(x)+\ldots+a_np_n(x)$$

$$i\neq 0 \quad \int_a^b P_i(x)P_0(x)\mu(x)dx=0$$

ומכאן שקיבלנו: $\int_a^b P_n(x)\mu(x)dx = a_0 \cdot \int_a^b \mu(x)dx = a_0 \cdot c$

את הפולינומים נמצא לפי תהליך גרהם שמידט:

$$f(x) \sim a_0 p_0 + a_1 p_1(x) + \ldots + a_n p_n(x) + R_n(x)$$

כאשר $p_i(x)$ אורתוגונאליים.

$$p_0(x) = 1 \quad \int_a^b P_k(x) P_0(x) \mu(x) dx = 0 \quad k \neq 0$$

$$\int_a^b f(x)\mu(x)dx = \int_a^b a_0 \mu(x) dx = a_0 c_0$$

אפשר למצוא את האינטגרל רק בעזרת a_0 ו- $\mu(x)$.

$$\int_a^b f(x)\mu(x)dx = \sum_{i=1}^{n} w_i f(x_i) \quad \left(f(x_i) = f_i\right)$$

נחשב רק עבור הפולינומים האורתוגונאליים (הבעיה: מהי הדרגה האלגברית של הנוסחה הנ"ל).

$$\begin{cases} 0 & j \neq 0 \\ c_0 & j = 0 \end{cases} = \int_a^b P_i(x)\mu(x)dx = \sum_{i=1}^{n} w_i P_j(x_i) \quad j = 0,1,\ldots,n-1$$

$$\begin{pmatrix} P_0(x_1) & P_0(x_2) & \ldots & P_0(x_n) \\ \vdots & & & \vdots \\ P_{n-1}(x_1) & P_{n-1}(x_2) & \ldots & P_{n-1}(x_n) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} c_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

נוסיף עוד שורת אפסים $\rightarrow P_n(x_1) \quad P_n(x_2) \quad \ldots \quad P_n(x_n)$

(x_i הם שורשים של פולינום אורתוגונאלי ממעלה n)