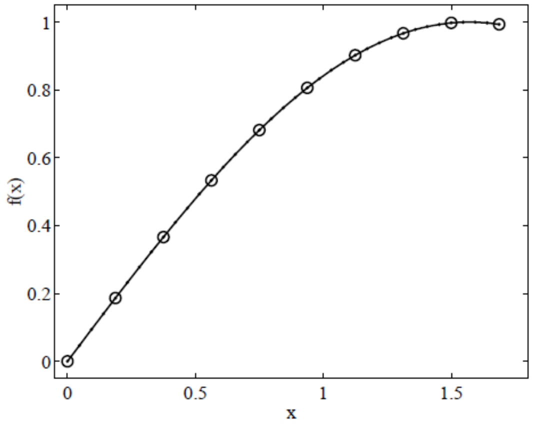
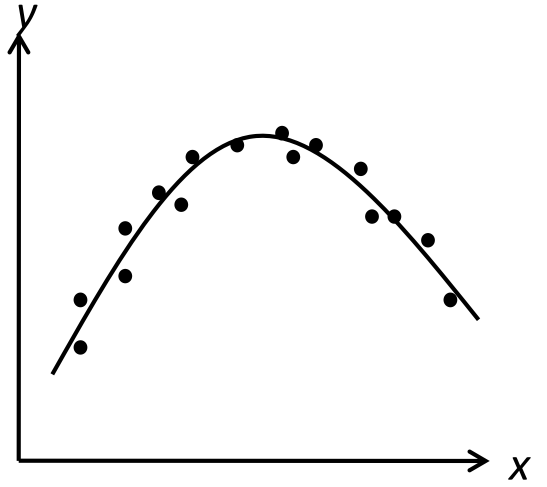
Aproximação de dados.

Dado um conjunto de dados tabulado $\{x_i, y_i\}$, encontrar a curva ou função que melhor representa os dados.



1. Interpolação.

$$y_i = f(x_i)$$



2. Aproximação dos mínimos quadrados

$$\text{minimize} \quad \sum_{i=1}^{n} (y_i - f(x_i))^2$$

Aproximação de dados. Interpolação polinomial.

Interpolação polinomial

Pretende-se encontrar o polinómio de grau N-1 que passa exactamente por N pontos $\{x_i,y_i\}$.

Este problema resume-se a resolver um sistema de N equações a N incógnitas:

$$\sum_{i=1}^{n} a_i x_k^{i-1} = f(x_k) \qquad k = 1, \cdots, n$$

À medida que N aumenta, a matriz desta equação torna-se mal condicionada dificultando a determinação exata dos coeficientes.

Exemplo:

Encontrar o polinómio para representar a pressão no intervalo [220,232]:

$T(^oF)$	$P(psia)$
220.0000	17.1860
224.0000	18.5560
228.0000	20.0150
232.0000	21.5670

Aproximação de dados. Interpolação polinomial.

Existência e unicidade.

Dados n+1 pontos:

$$\left(x_0, f(x_0)\right), \left(x_1, f(x_1)\right), \ldots., \left(x_n, f(x_n)\right)$$

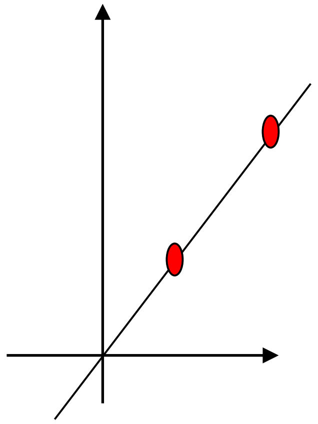
Assumindo que $x_0, x_1, \ldots, x_n$ são distintos

Teorema:

Existe um único polinómio ***$f_n(x)$*** de ordem ≤ n:

$$f_n(x_i) = f(x_i) \qquad for \quad i = 0,1,\ldots,n$$

Aproximação de dados. Interpolação linear.



Dados dois pontos quaiquer, $\left(x_0, f(x_0)\right), \left(x_1, f(x_1)\right)$

A recta que interpola esses dois pontos é:

$$f_1(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}\left(x - x_0\right)$$

Exemplo :

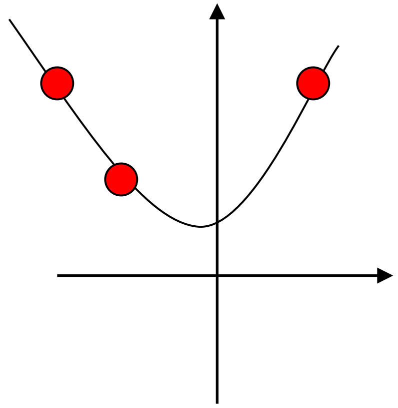
Qual o polinómio que interpola (1,2) and (2,4).

$$f_1(x) = 2 + \frac{4-2}{2-1}\left(x-1\right) = 2x$$

Aproximação de dados. Interpolação quadrática.

- Dados três pontos quaisquer: $(x_0, f(x_0))$, $(x_1, f(x_1))$, e $(x_2, f(x_2))$
- O polinómio que interpolador na forma de Newton é:

$$\boxed{f_2(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1)}$$



Onde:

$$b_0 = f(x_0)$$

$$b_1 = f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

$$b_2 = f[x_0, x_1, x_2] = \frac{\dfrac{f(x_2) - f(x_1)}{x_2 - x_1} - \dfrac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_0}$$

Aproximação de dados. Interpolação ordem n.

Dados n+1 pontos, o polinómio interpolador de Newton é:

$$f_n(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1) + \ldots + b_n(x - x_0)\ldots(x - x_{n-1})$$

$$b_0 = f(x_0)$$

$$b_1 = f[x_0, x_1]$$

....

$$b_n = f[x_0, x_1, \ldots, x_n]$$

$$f_n(x) = \sum_{i=0}^{n} \left\{ F[x_0, x_1, \ldots, x_i] \prod_{j=0}^{i-1} \left(x - x_j\right) \right\}$$

Interpolação polinomial ordem n. Diferenças divididas.

$$f[x_k] = f(x_k)$$ **DD Ordem 0**.

$$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0}$$ **DD Ordem 1**.

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$ **DD Ordem 2**.

....

$$f[x_0, x_1, ..., x_k] = \frac{f[x_1, x_2, ..., x_k] - f[x_0, x_1, ..., x_{k-1}]}{x_k - x_0}$$ **DD Ordem k**.

x	f[ ]	f[,]	f[, ,]	f[, , ,]
x0	f[x0]	f[x0,x1]	f[x0,x1,x2]	f[x0,x1,x2,x3]
x1	f[x1]	f[x1,x2]	f[x1,x2,x3]	
x2	f[x2]	f[x2,x3]		
x3	f[x3]			

Interpolação polynomial ordem n. Diferenças divididas.

x_i	$f(x_i)$
0	-5
1	-3
-1	-15

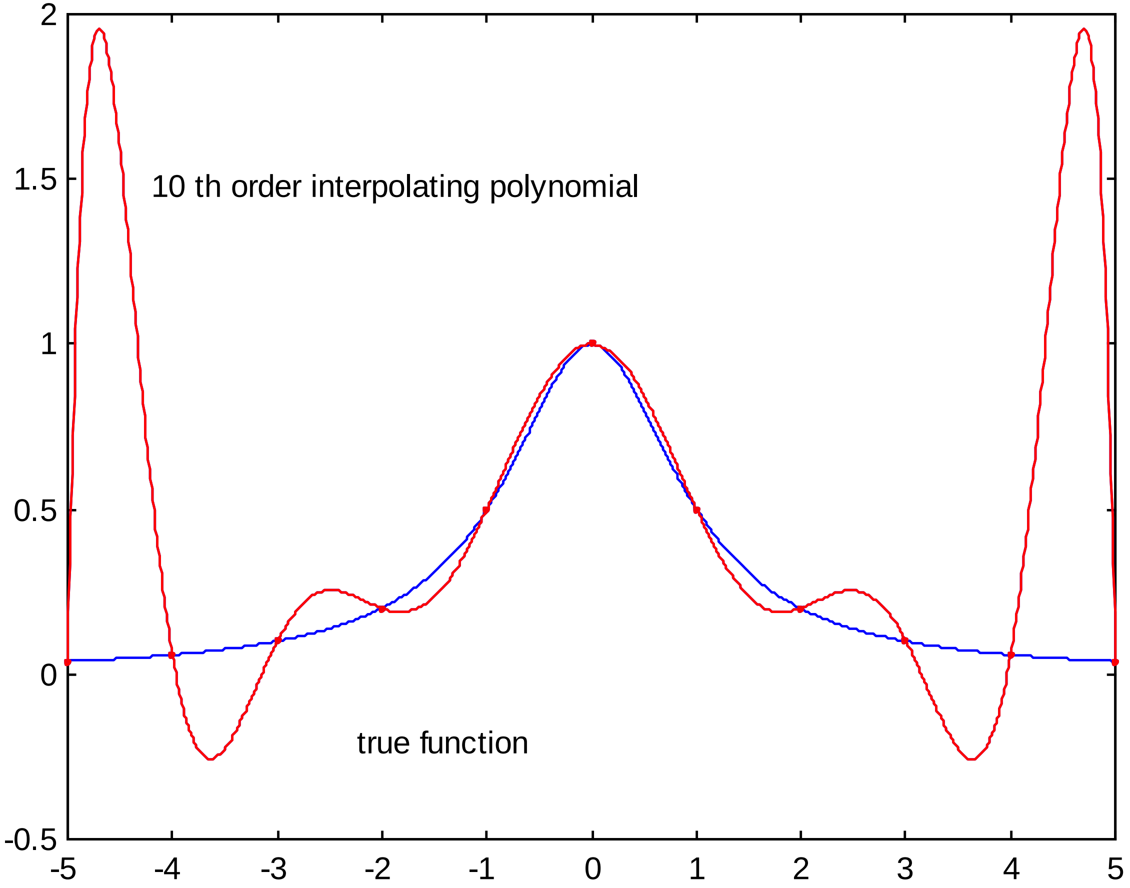
Qual o polinómio interpolador de Newton para os dados da tabela?

x	F[ ]	F[,]	F[, ,]
0	-5	2	-4
1	-3	6	
-1	-15		

$$f_2(x) = -5 + 2(x-0) - 4(x-0)(x-1)$$

$$f_2(x) = F[x_0] + F[x_0, x_1](x-x_0) + F[x_0, x_1, x_2](x-x_0)(x-x_1)$$

Interpolação Polinomial ordem 10.



Aproximação de dados. Método dos Minimos quadrados.

Método Geral.

Em geral pretende-se encontrar a função que melhor se ajusta a um conjunto de N pontos $\{x_i,y_i\}$.

$$y_k \approx P_{M-1}(x_k) = \sum_{i=1}^{M} c_i \phi_i(x_k) \qquad \Leftrightarrow \qquad A.c = y$$

Selecionando os coeficientes c_i por forma minimizar o erro quadrático:

$$\varepsilon = \sum_{k=1}^{N} (P_{M-1}(x_k) - y_k)^2 \qquad \Leftrightarrow \qquad \|d\|_2^2 = \|y - A.c\|_2^2$$

$$\frac{\partial \varepsilon}{\partial c_i} = 0 \qquad i = 1, \dots, M$$

Aproximação de dados. Método dos Minimos quadrados.

Desenvolvendo a expressão

$$\|d\|^2 = \|y - Ac\|^2 = (y - Ac)^T (y - Ac) =$$
$$= y^T y - 2y^T Ac - c^T A^T AC$$

Derivando em ordem a c e igualando a zero

$$-2y^T A + 2c^T A^T A = 0$$

$$\Leftrightarrow \boxed{A^T Ac = A^T y}$$

Se a matriz $A^T A$ possuir inversa (se as colunas de A forem lin. indep.) então

$$\boxed{c = (A^T A)^{-1} A^T y}$$

Os coeficientes c_i da função de ajuste podem assim ser obtidos resolvendo o sistema de equações sobredeterminado (A.c=y) usando a pseudo inversa de A!

Método dos Minimos quadrados Linear.

Em alternativa, pode-se resolver o sistema de M equações normais resultante das derivadas parciais de ε:

$$\begin{bmatrix} \sum_{i=1}^{N} \Phi_1(x_i)\Phi_1(x_i) & \sum_{i=1}^{N} \Phi_1(x_i)\Phi_2(x_i) \cdots & \sum_{i=1}^{N} \Phi_1(x_i)\Phi_{M-1}(x_i) & \sum_{i=1}^{N} \Phi_1(x_i)\Phi_M(x_i) \\ & & \vdots & \\ & & \vdots & \\ \sum_{i=1}^{N} \Phi_M(x_i)\Phi_1(x_i) & \sum_{i=1}^{N} \Phi_M(x_i)\Phi_2(x_i) \cdots & \sum_{i=1}^{N} \Phi_M(x_i)\Phi_{M-1}(x_i) & \sum_{i=1}^{N} \Phi_M(x_i)\Phi_M(x_i) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} y_i \Phi_1(x_i) \\ \vdots \\ \vdots \\ \sum_{i=1}^{N} y_i \Phi_M(x_i) \end{bmatrix}$$

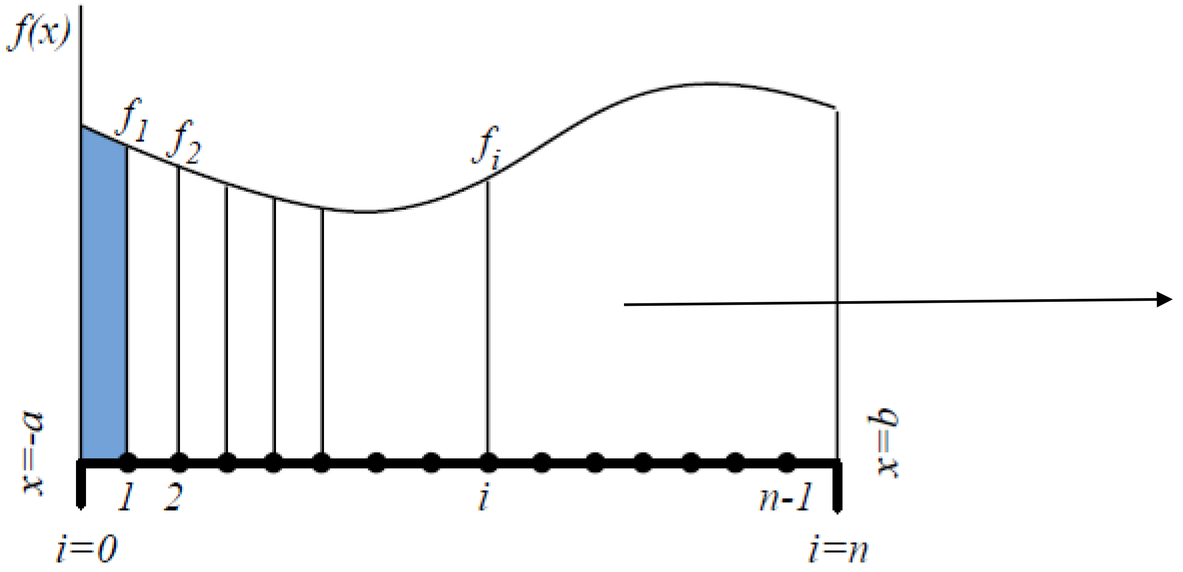
Método dos Minimos quadrados Linear.

Por exemplo, no caso da regressão linear temos :

$$f(x,t) = a + bx + ct, \quad \Phi(a,b,c) = \sum_{i=1}^{n} \left(a + bx_i + ct_i - y_i\right)^2$$

$$\begin{aligned}
\frac{\partial \Phi(a,b,c)}{\partial a} &= 2\sum_{i=1}^{n}\left(a + bx_i + ct_i - y_i\right) = 0 \\
\frac{\partial \Phi(a,b,c)}{\partial b} &= 2\sum_{i=1}^{n}\left(a + bx_i + ct_i - y_i\right)x_i = 0 \\
\frac{\partial \Phi(a,b,c)}{\partial c} &= 2\sum_{i=1}^{n}\left(a + bx_i + ct_i - y_i\right)t_i = 0
\end{aligned}
\Leftrightarrow
\left\{\begin{aligned}
&a\,n + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} t_i = \sum_{i=1}^{n} y_i \\
&a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} (x_i)^2 + c\sum_{i=1}^{n} (x_i\ t_i) = \sum_{i=1}^{n} (x_i\ y_i) \\
&a\sum_{i=1}^{n} t_i + b\sum_{i=1}^{n} (x_i\ t_i) + c\sum_{i=1}^{n} (t_i)^2 = \sum_{i=1}^{n} (t_i\ y_i)
\end{aligned}\right.$$

Integração Numérica.



$$Area = \int_a^b f(x)dx$$

Divide-se o intervalo [a,b] em n subintervalos e aproxima-se a função por um polinómio de Newton-Cotes de grau n:

$$f(x) = P_n(x_0 + \alpha h) + \mathcal{O}(h^{n+1})$$

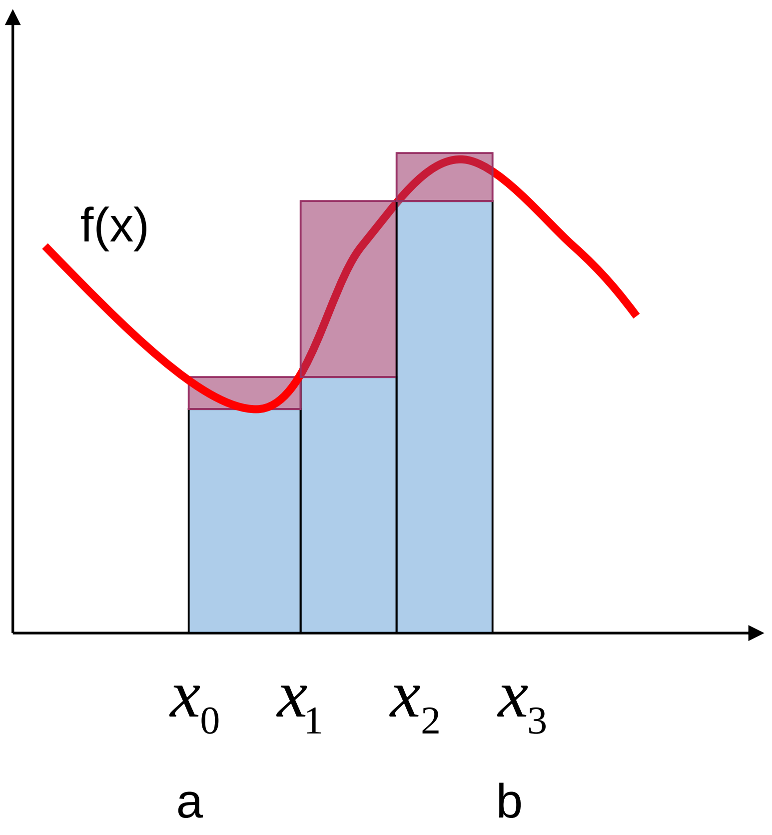
Regras do rectângulo e do ponto médio.

A **regra do retângulo** consiste em aproximar a função por um polinómio de grau 0 em cada subintervalo.

$$\int_a^b f(x)dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)dx = h \sum_{i=1}^{n} f_{i-1} + \mathcal{O}(h)$$

Ou

$$\int_a^b f(x)dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)dx = h \sum_{i=1}^{n} f_{i-1} + \mathcal{O}(h)$$



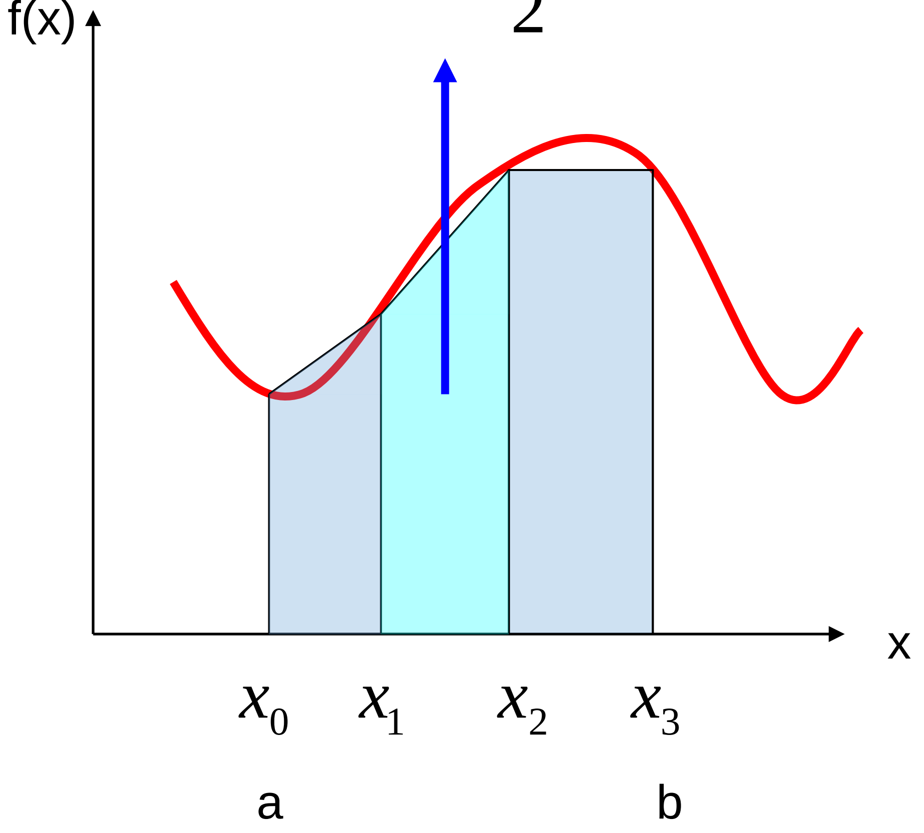
Na **regra do ponto médio** faz-se avaliação da função no centro do subintervalo.

$$\int_a^b f(x)dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)dx = h \sum_{i=1}^{n} f\left(\frac{x_i + x_{i-1}}{2}\right) + \mathcal{O}(h^2)$$

Regra do Trapézio.

A **regra do trapézio** consiste em aproximar a função por um polinómio de grau 1 em cada subintervalo.

$$Area = \frac{f(x_2) + f(x_1)}{2}(x_2 - x_1)$$



$$\int_a^b f(x)dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)dx$$

$$= \sum_{i=1}^{n} \frac{h}{2}[f_{i-1} + f_i] + \mathcal{O}(h^2)$$

Regra do Trapézio. Exemplo.

Dados os valores da velocidade de um objecto, obter uma estimative da distância percorrida no interval de tempo [0,3].

Time (s)	0.0	1.0	2.0	3.0
Velocity (m/s)	0.0	10	12	14

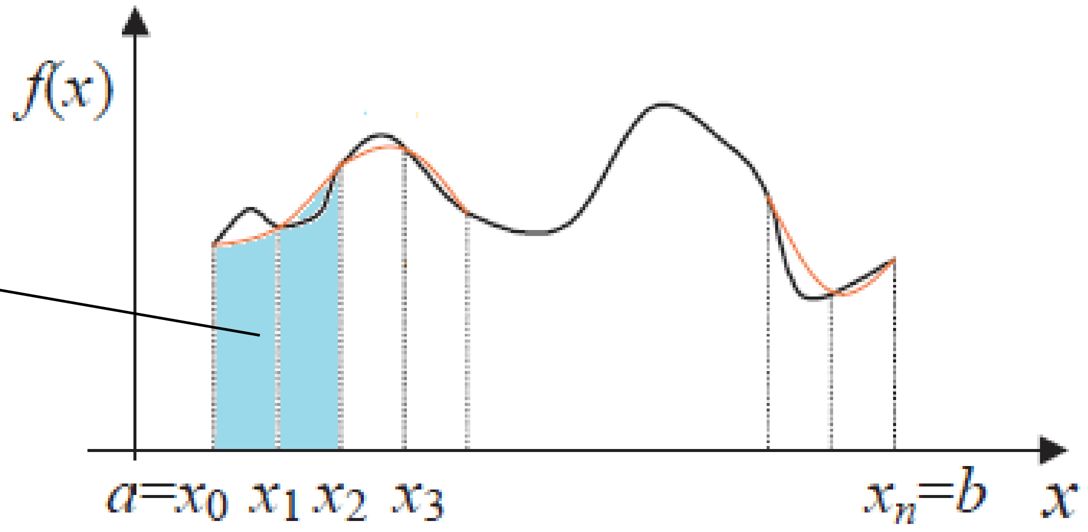
$$Distância = \int_0^3 V(t)dt = \frac{1}{2}\big((10+0)+(12+10)+(14+12)\big)$$

= 5 + 11 + 13 = 29 m

Regra de Simpson.

A **regra do Simpson** consiste em aproximar a função por um polinómio de grau 2 em cada 2 subintervalos.

$$\int_{x_0}^{x_2} f(x)\,dx \approx \frac{x_2-x_0}{6}\left[f(x_0)+4f(x_1)+f(x_2)\right]$$



$$\int_a^b f(x)dx = \frac{h}{3}\left[f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \cdots + f_n\right] + \quad O(h^4)$$

Extrapolação de Romberg.

A ideia é estimar o erro de truncatura avaliando o integral em 2 grelhas com passos diferentes h_1 e h_2 :

$$I = I(h_1) + E(h_1) = I(h_2) + E(h_2)$$

$$\frac{E(h_1)}{E(h_2)} = \frac{h_1^2}{h_2^2}$$ (Na regra do trapézio)

Combinando as equações temos:

$$E(h_2) = \frac{I(h_1) - I(h_2)}{[1 - (h_1/h_2)^2]}$$

$$I = I(h_2) + \frac{1}{[(h_1/h_2)^2 - 1]}[I(h_2) - I(h_1)] + \mathcal{O}(h^4)$$

Extrapolação de Romberg.

Aplicando esta regra dividindo sucessivamente o intervalo a metade h, h/2, …:

1ª coluna obtida com método ponto médio, trapézio, Simpson

$$I_{j,k} = \frac{4^k I_{j,k-1} - I_{j-1,k-1}}{4^k - 1}$$

I(0,0)			
I(1,0)	I(1,1)		
I(2,0)	I(2,1)	I(2,2)	
I(3,0)	I(3,1)	I(3,2)	I(3,3)

Outros elementos obtidos com método de Romberg

Erro Trapez. $O(h^2)$ $O(h^4)$ $O(h^6)$ $O(h^8)$

Teorema

$$\int_a^b f(x)dx = \mathrm{I}(n,m) + O(h^{2m+2})$$

Extrapolação de Romberg.

Calcule: $\int_0^1 x^2 dx$

0.5		
3/8	1/3	
11/32	1/3	1/3

$$h = 1,\ R(0,0) = \frac{b-a}{2}\left[f(a) + f(b)\right] = \frac{1}{2}\left[0+1\right] = 0.5$$

$$h = \frac{1}{2},\ R(1,0) = \frac{1}{2}R(0,0) + h(f(a+h)) = \frac{1}{2}\left(\frac{1}{2}\right) + \frac{1}{2}\left(\frac{1}{4}\right) = \frac{3}{8}$$

$$R(1,1) = \frac{1}{4^1 - 1}\left[4 \times R(1,0) - R(0,0)\right] = \frac{1}{3}\left[4 \times \frac{3}{8} - \frac{1}{2}\right] = \frac{1}{3}$$

$$h = \frac{1}{4},\ R(2,0) = \frac{1}{2}R(1,0) + h(f(a+h) + f(a+3h)) = \frac{1}{2}\left(\frac{3}{8}\right) + \frac{1}{4}\left(\frac{1}{16} + \frac{9}{16}\right) = \frac{11}{32}$$

$$R(2,1) = \frac{1}{3}\left[4 \times R(2,0) - R(1,0)\right] = \frac{1}{3}\left[4 \times \frac{11}{32} - \frac{3}{8}\right] = \frac{1}{3}$$

$$R(2,2) = \frac{1}{4^2 - 1}\left[4^2 \times R(2,1) - R(1,1)\right] = \frac{1}{15}\left[\frac{16}{3} - \frac{1}{3}\right] = \frac{1}{3}$$

Regra da quadratura de Gauss.

A Quadratura de Gauss escolhe os pontos para calculo de uma forma otima, em vez de igualmente espaçada como acontece nas regras anteriores (rectângulo, Simpson).

Os pontos $x_1; x_2; \dots; x_n$ no intervalo [a; b] e os coefientes $w_1; w_2; \dots; w_n$ são escolhidos para minimizar o erro esperado na aproximação, fornecendo 2n parâmetros a determinar.

$$\int_{-1}^{1} f(x)\,dx \approx \sum_{i=1}^{n} w_i f(x_i)$$

Note que temos n coeficientes w_i e n pontos x_i para determinar. O problema de encontrar os pesos e abscissas é equivalente a um sistema não linear com equações e incógnitas. Este problema sempre tem solução e a solução é única se $x_1 < x_2 < \dots < x_n$.

A regra é precisa para polinómios de grau 2n-1.

Regra da quadratura de Gauss.

Os nós x_i são dados pelos zeros do polinômio de Legendre, $P_n(x)$:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x) , \qquad P_0(x) = 1 , \quad P_1(x) = x.$$

Os pesos são dados por :

$$w_i = \frac{2}{\left(1 - x_i^2\right) \left[P'_n(x_i)\right]^2}.$$

Para aproximar a integral de f(x) no intervalo [a,b] devemos fazer a mudança de variável

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2} \sum_{i=1}^{n} w_i f\left(\frac{b-a}{2}\xi_i + \frac{a+b}{2}\right).$$

Regra da quadratura de Gauss.

Number of points, n	Points, x_i		Weights, w_i	
1	0		2	
2	$\pm\frac{1}{\sqrt{3}}$	±0.57735...	1	
3	0		$\frac{8}{9}$	0.888889...
	$\pm\sqrt{\frac{3}{5}}$	±0.774597...	$\frac{5}{9}$	0.555556...
4	$\pm\sqrt{\frac{3}{7}-\frac{2}{7}\sqrt{\frac{6}{5}}}$	±0.339981...	$\frac{18+\sqrt{30}}{36}$	0.652145...
	$\pm\sqrt{\frac{3}{7}+\frac{2}{7}\sqrt{\frac{6}{5}}}$	±0.861136...	$\frac{18-\sqrt{30}}{36}$	0.347855...
5	0		$\frac{128}{225}$	0.568889...
	$\pm\frac{1}{3}\sqrt{5-2\sqrt{\frac{10}{7}}}$	±0.538469...	$\frac{322+13\sqrt{70}}{900}$	0.478629...
	$\pm\frac{1}{3}\sqrt{5+2\sqrt{\frac{10}{7}}}$	±0.90618...	$\frac{322-13\sqrt{70}}{900}$	0.236927...

Integrais impróprios.

Se os limites de integração são infinitos ou a função de integração não é limitada para um dos limites, as regras de integração do trapézio, Simpson não podem ser usadas.

- Mudança de variável

$$\int_a^b f(x)dx = \int_{\frac{1}{b}}^{\frac{1}{a}} \frac{1}{t^2} f\left(\frac{1}{t}\right)dt, \quad \text{(assuming } ab > 0)$$

ou

$$\int_0^\infty f(x)\,dx = \int_0^1 \frac{f(-\ln t)}{t}\,dt$$

$$\textit{Exemplo}: \quad \int_1^\infty \frac{1}{x^2}\,dx = \int_0^1 \frac{1}{t^2}\left(\frac{1}{\left(\frac{1}{t}\right)^2}\right)dt$$

- Truncar o integral

$$I = \int_a^c f(x)\,dx + \cancel{\int_c^\infty f(x)\,dx}$$

- Usar uma regra de integração que evite os extremos.