Aproximación y modelos

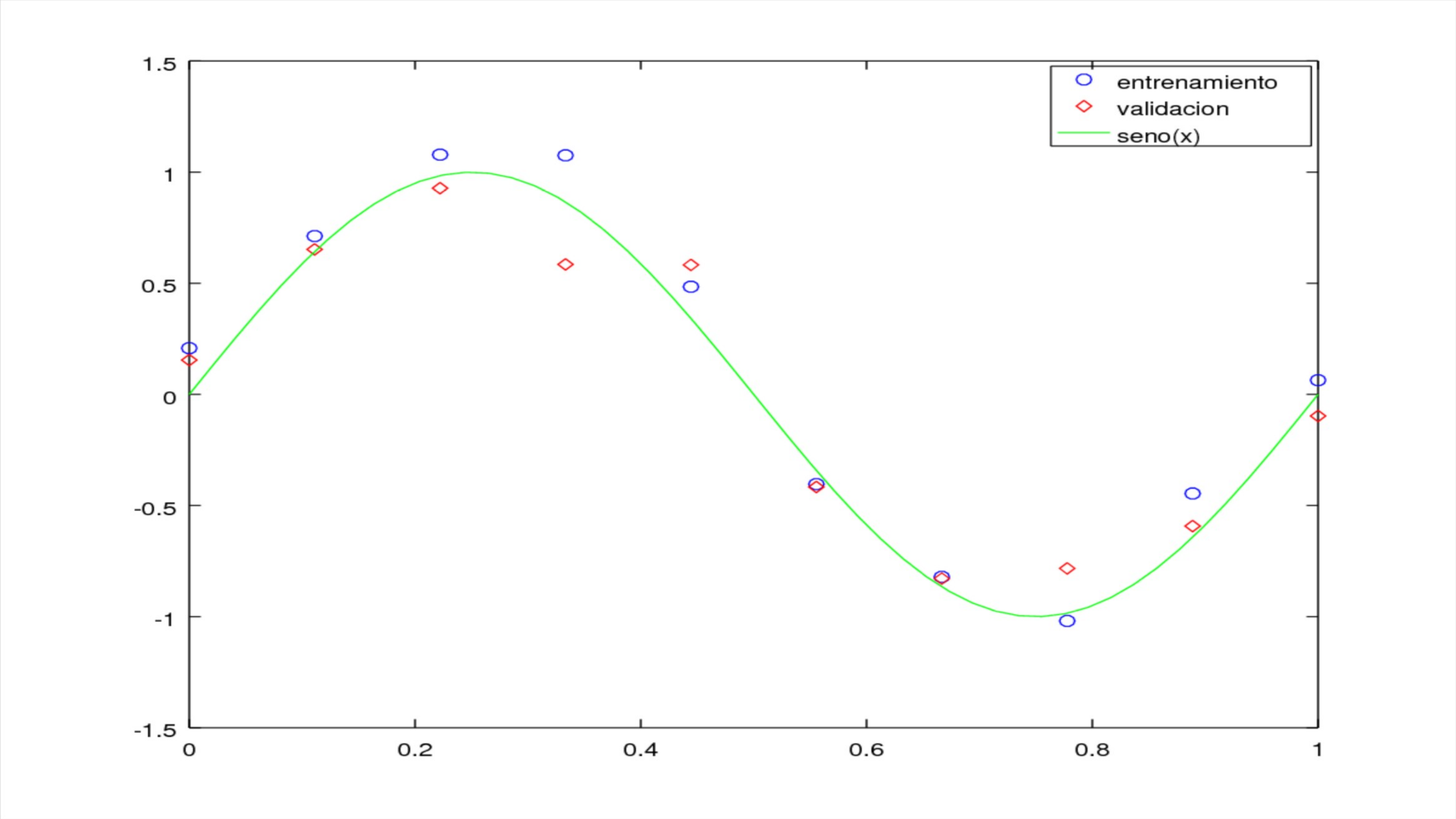
Aproximación polinomial, regularización y modelos

Miguelangel Fraga Aguilar

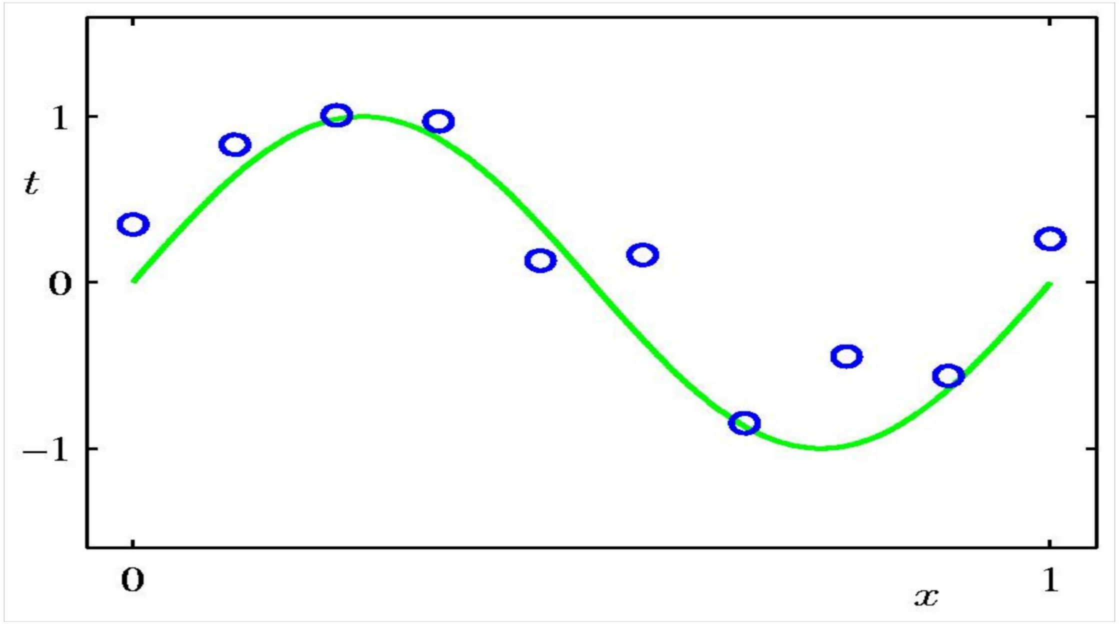
Problema de ejemplo

```
%archivo para generar datos de entrenamiento y
% prueba para regresion plinomial
x=linspace(0,1,10)';
entrenamiento=sin(2*pi*x)+0.2*randn(size(x));
validacion=sin(2*pi*x)+0.2*randn(size(x));
%guarda los datos en un archivo .mat
save seno_ruido.mat
%muesra y grafica los datos junto con la funcion seno
disp('Datos generados');
disp(x);
disp(entrenamiento);
disp(validacion);
xg=linspace(0,1,50)';
plot(x,entrenamiento,'ob',x,validacion,'dr',xg,sin(2*pi*xg),'-g');
legend('entrenamiento','validacion','seno(x)');

```

Problema de ejemplo (2)

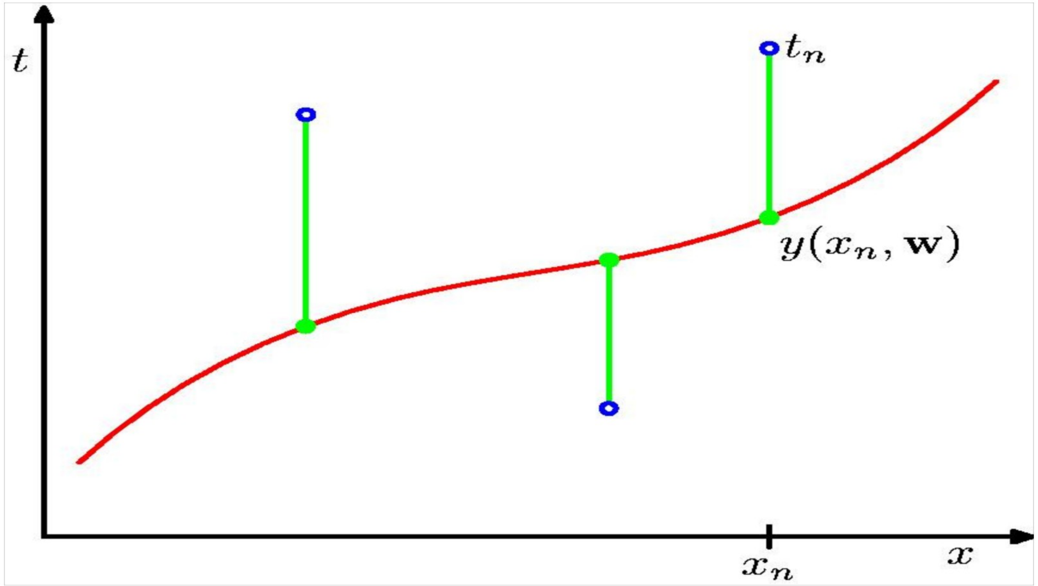


Ajuste de curvas polinomiales



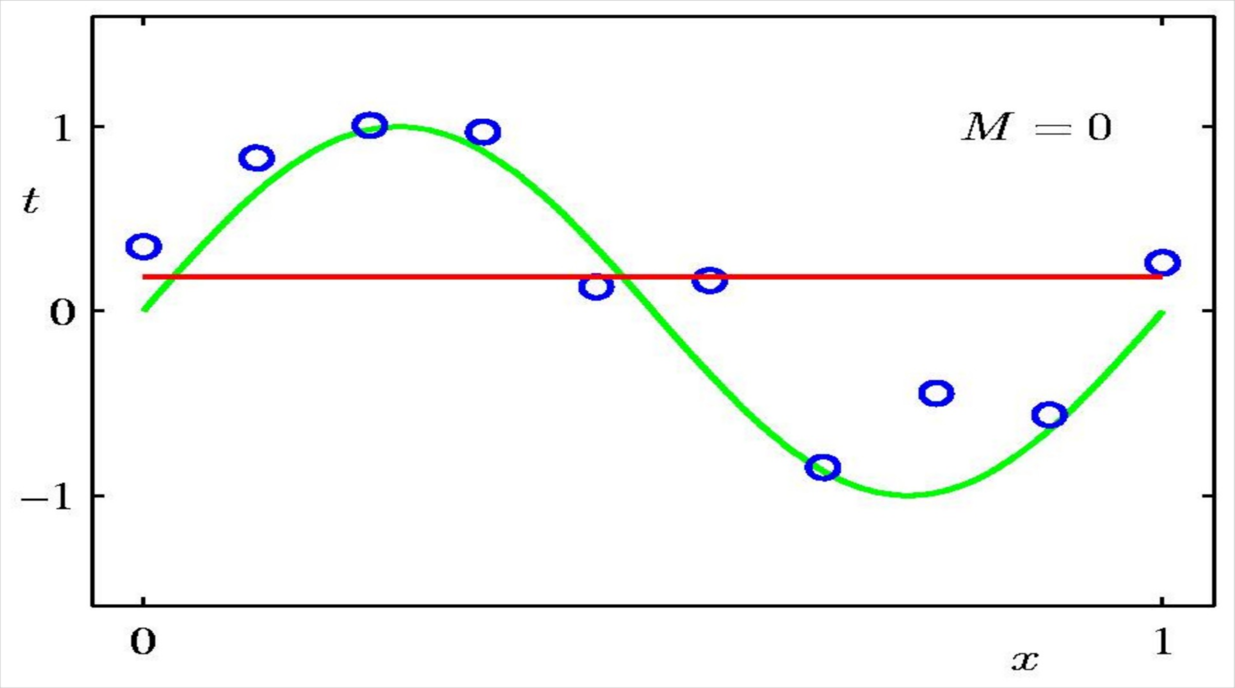
$$y(x, \mathbf{w}) = w_0 + w_1 x + w_2 x^2 + \ldots + w_M x^M = \sum_{j=0}^{M} w_j x^j$$

Función de suma de los cuadrados de los errores

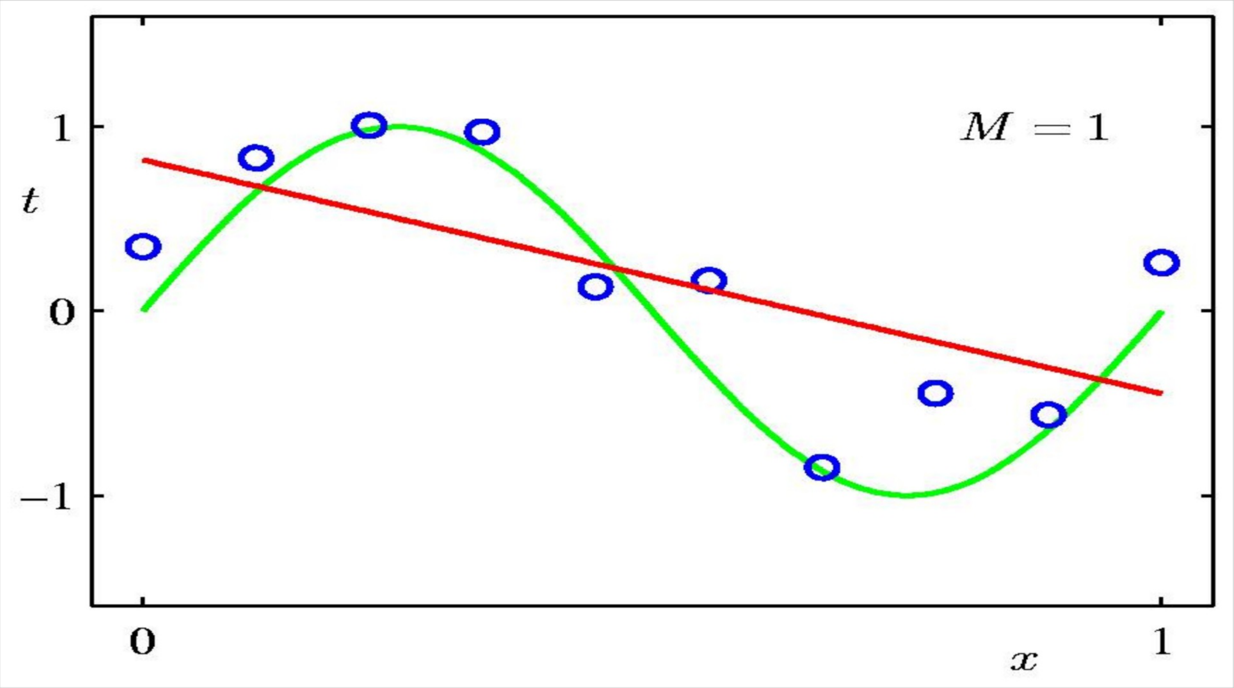


$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} \{y(x_n, \mathbf{w}) - t_n\}^2$$

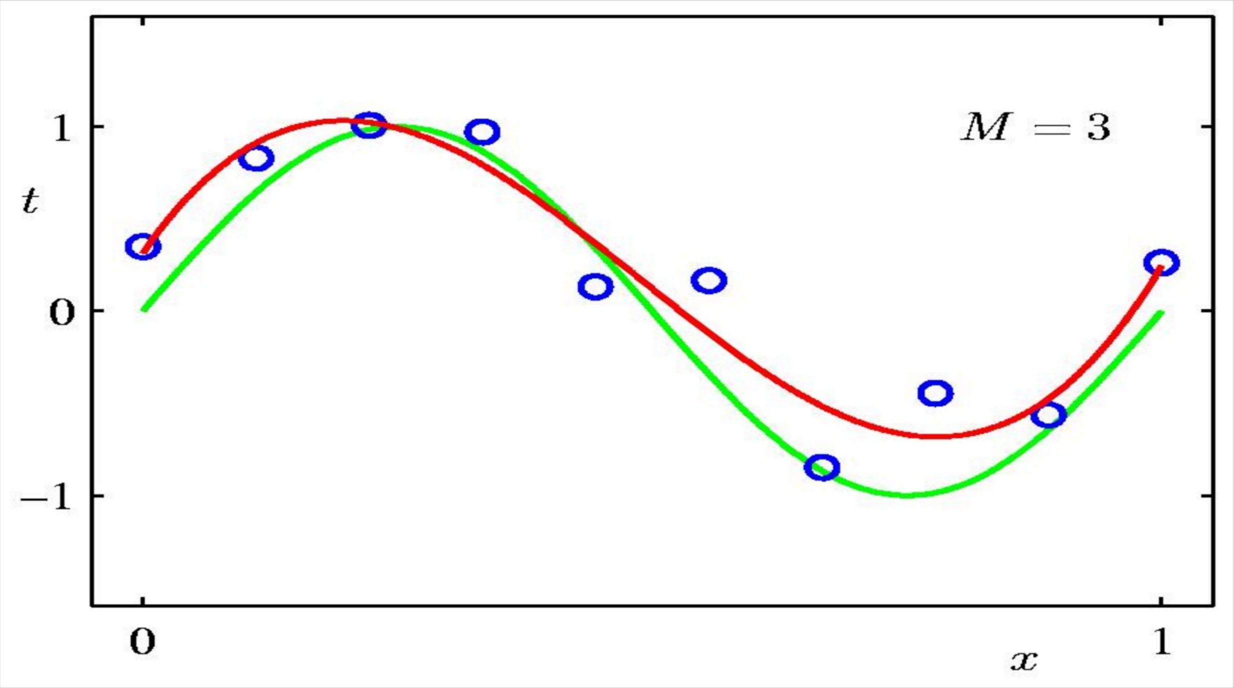
Polinomio de grado 0



Polinomio de grado 1



Polinomio de grado 3



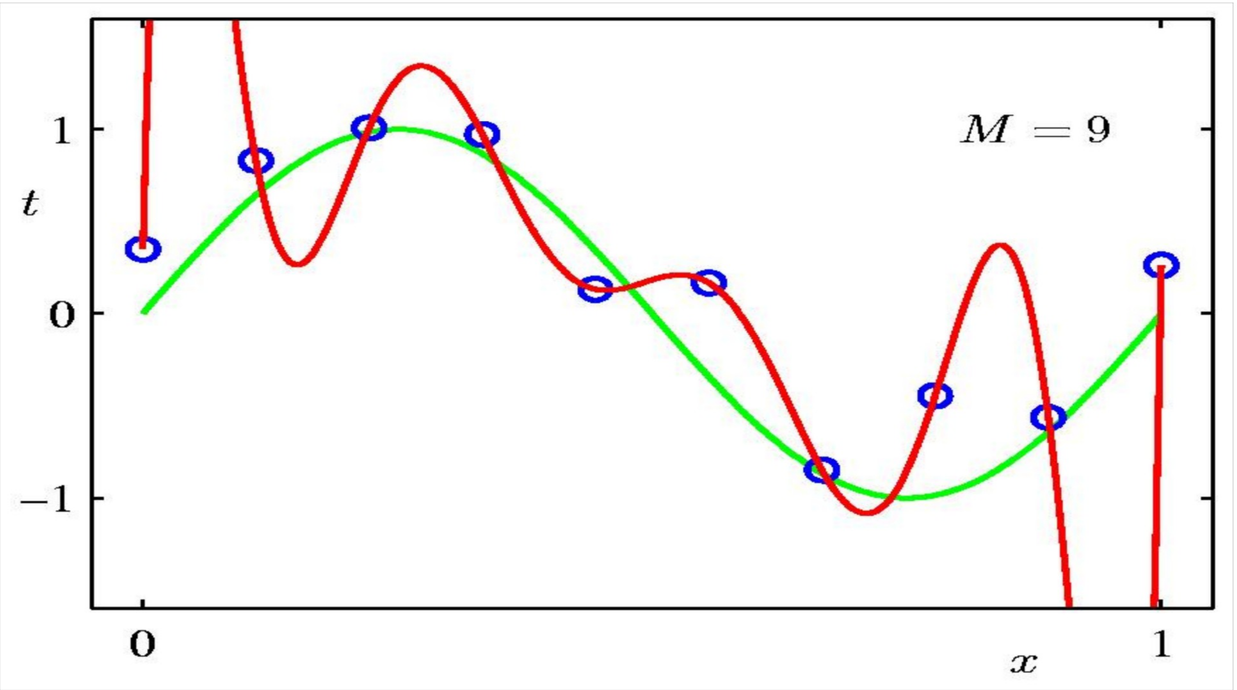
Guión de octave

```
%aproxima un conjunto de datos con un polinomio de grado m

%lee los datos del archivo
load seno_ruido.mat

%grado del polinomio de aproximacion
m=9;
n=length(x);
%Construye la matriz de disenio
A=zeros(n,m+1);
for indice=1:m+1
  A(:,indice)=x.^(indice-1);
end
%calcula los pesos de los polinomios
w=A\entrenamiento;
disp('pesos');
disp(w);
```

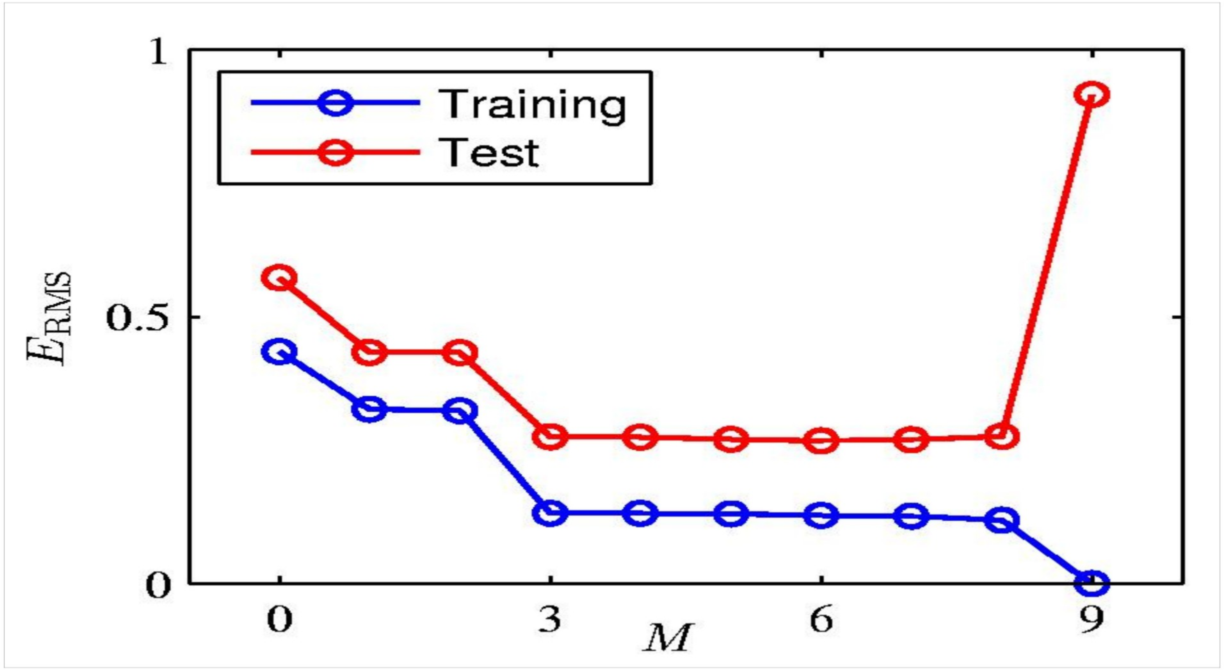
Guión de octave (2)

```
errores=A*w-entrenamiento;
fprintf('Error rms entrenamiento=%f\n',sqrt(dot(errores,errores)/n));

errores=A*w-validacion;
fprintf('Error rms validacion=%f\n',sqrt(dot(errores,errores)/n));

%grafica el polinomio resultante
xg=linspace(0,1,50)';
%Construye la matriz de potencias de xg
Ag=zeros(50,m+1);
for indice=1:m+1
  Ag(:,indice)=xg.^(indice-1);
end

plot(x,entrenamiento,'ob',x,validacion,'dr',xg,Ag*w,'-g',xg,sin(2*pi*xg),'-c');
legend('entrenamiento','validacion','y(x)','sin(x)');
```

Polinomio de grado 9



Sobre-entrenamiento



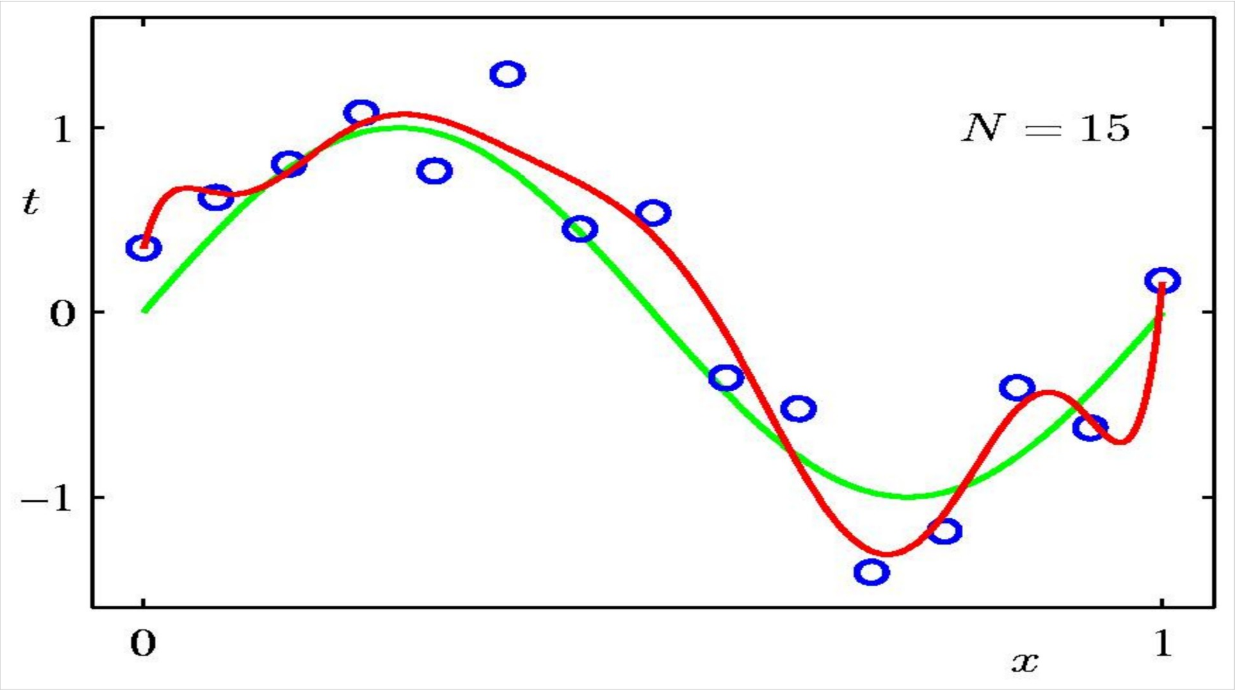
Root-Mean-Square (RMS) Error: $E_{\mathrm{RMS}} = \sqrt{2E(\mathbf{w}^\star)/N}$

Coeficientes de los polinomios

	$M = 0$	$M = 1$	$M = 3$	$M = 9$
$w_0^\star$	0.19	0.82	0.31	0.35
$w_1^\star$		-1.27	7.99	232.37
$w_2^\star$			-25.43	-5321.83
$w_3^\star$			17.37	48568.31
$w_4^\star$				-231639.30
$w_5^\star$				640042.26
$w_6^\star$				-1061800.52
$w_7^\star$				1042400.18
$w_8^\star$				-557682.99
$w_9^\star$				125201.43

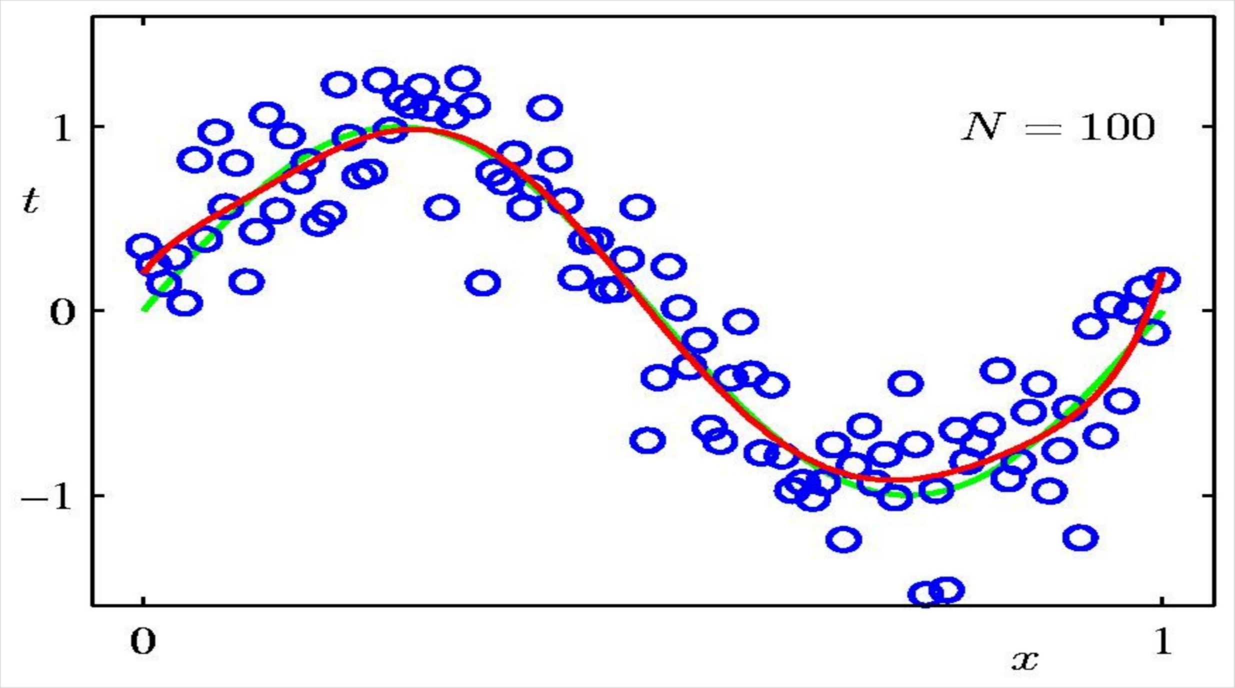
Tamaño del conjunto de datos $N = 15$

Polinomio de grado 9



Tamaño del conjunto de datos $N = 100$

Polinomio de grado 9



Aproximación y modelos

Regularización

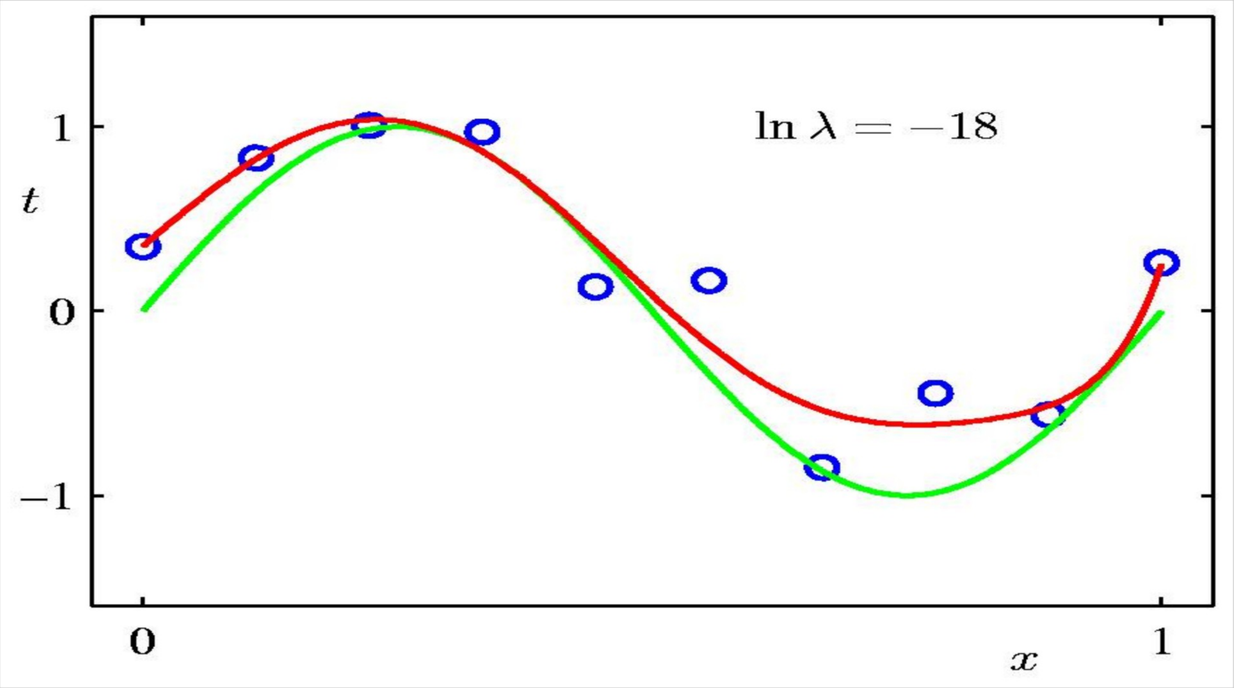
Miguelangel Fraga Aguilar

Regularización

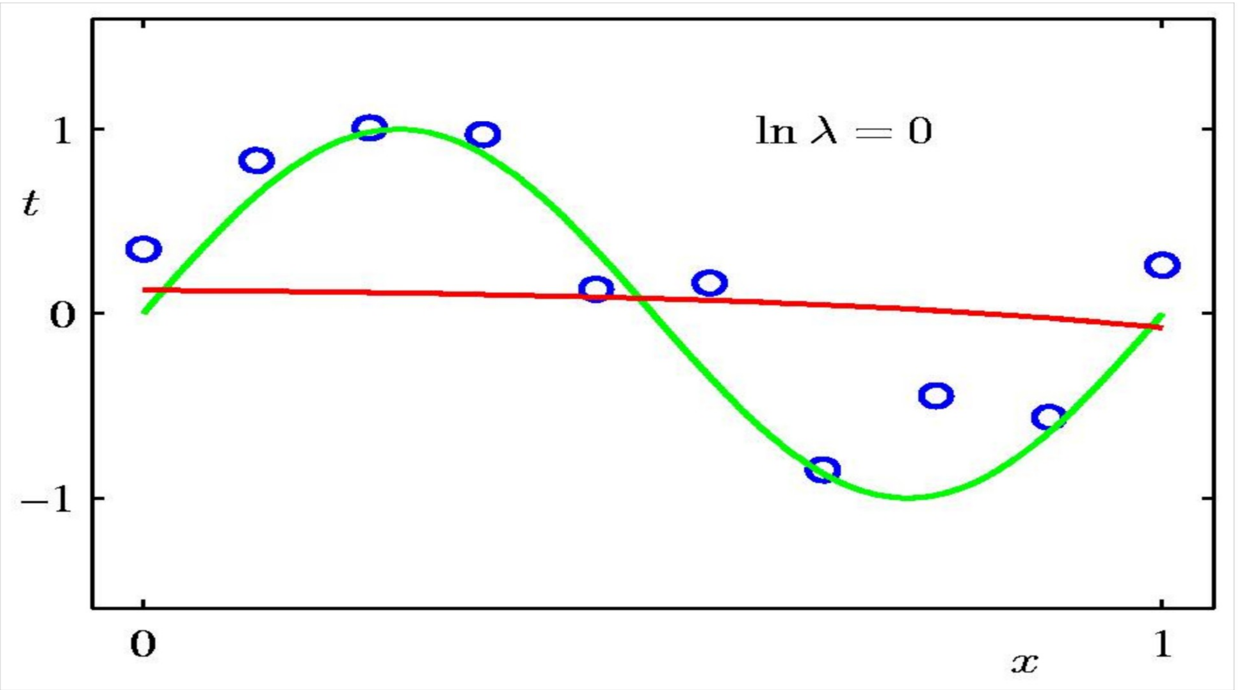
Penaliza los valores de coeficientes grandes

$$\widetilde{E}(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2}\|\mathbf{w}\|^2$$

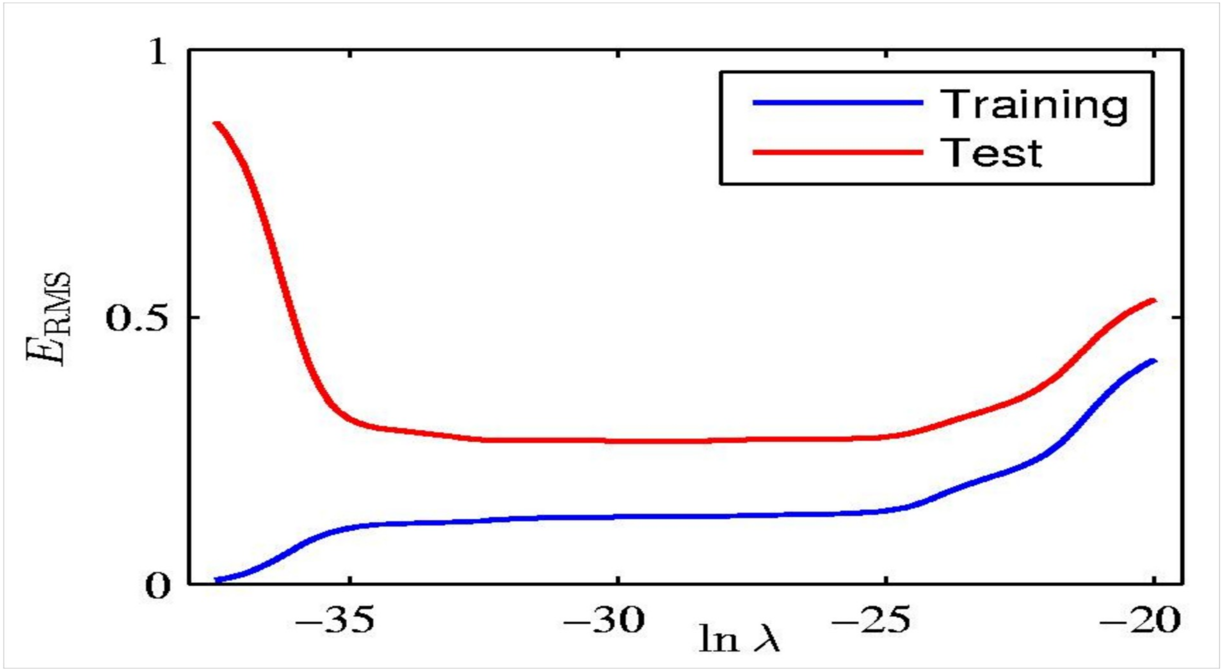
Regularización: $\ln \lambda = -18$



Regularización: $\ln \lambda = 0$



Regularización: E_{RMS} vs. $\ln \lambda$



Coeficientes Polinomiales

	$\ln \lambda = -\infty$	$\ln \lambda = -18$	$\ln \lambda = 0$
$w_0^\star$	0.35	0.35	0.13
$w_1^\star$	232.37	4.74	-0.05
$w_2^\star$	-5321.83	-0.77	-0.06
$w_3^\star$	48568.31	-31.97	-0.05
$w_4^\star$	-231639.30	-3.89	-0.03
$w_5^\star$	640042.26	55.28	-0.02
$w_6^\star$	-1061800.52	41.32	-0.01
$w_7^\star$	1042400.18	-45.95	-0.00
$w_8^\star$	-557682.99	-91.53	0.00
$w_9^\star$	125201.43	72.68	0.01

Aproximación y modelos

Otras funciones base

Miguelangel Fraga Aguilar

Modelos de funciones base

Generalizando

$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=0}^{M-1} w_j \phi_j(\mathbf{x}) = \mathbf{w}^{\mathrm{T}} \boldsymbol{\phi}(\mathbf{x})$$

donde $\varphi_j(x)$ son conocidas como funciones base.

Típicamente, $\varphi_0(x) = 1$, para que w_0 actúe como bias.

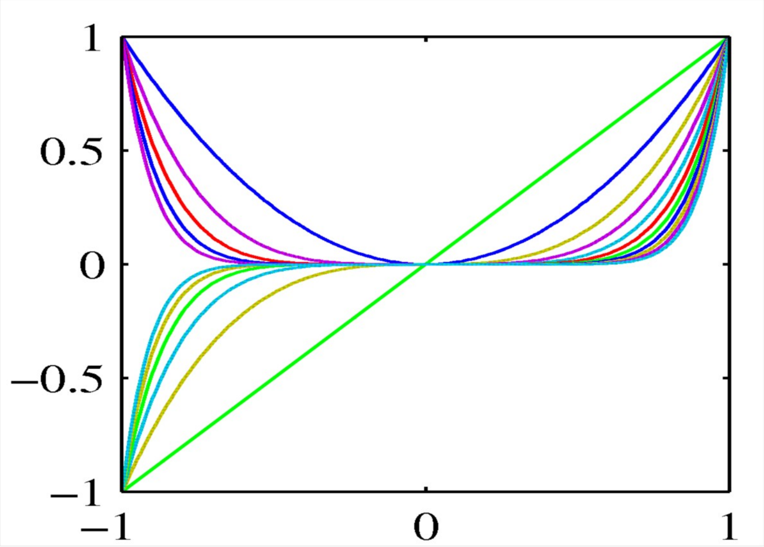
En el caso más simple, se usan funciones base lineales: $\varphi_d(x) = x_d$.

Modelos de funciones base (2)

Funciones base polinómicas:

$$\phi_j(x) = x^j.$$

Son globales, un pequeño cambio en x afecta a todas las funciones base

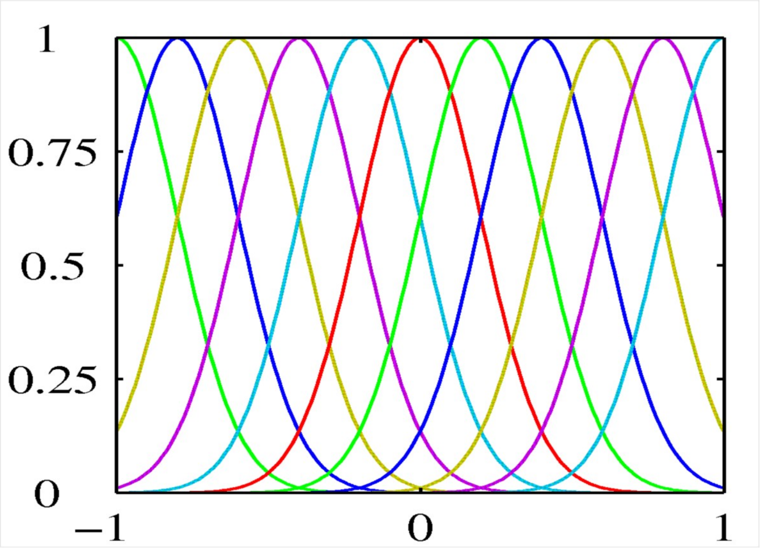


Modelos de funciones base (3)

Funciones base Gausianas:

$$\phi_j(x) = \exp\left\{-\frac{(x-\mu_j)^2}{2s^2}\right\}$$

son locales; un pequeño cambio en x solo afecta la funciones base cercanas. μ_j y s controlan la posición y la escala (ancho).



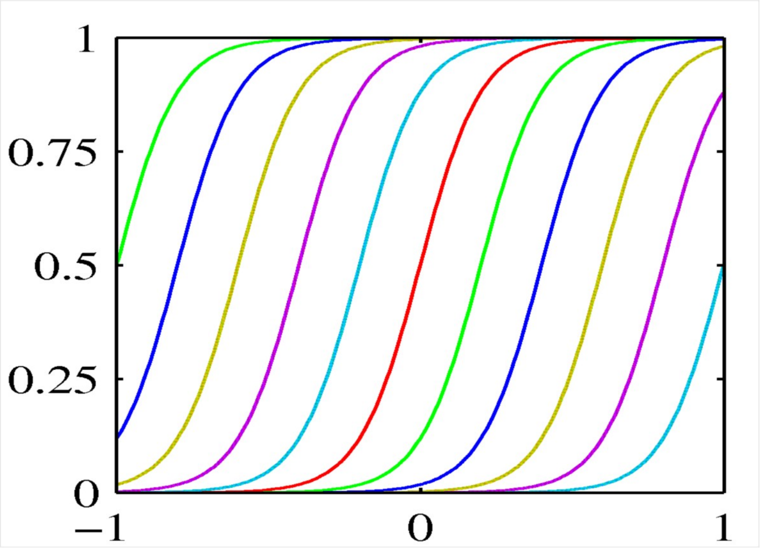
Modelos de funciones base (4)

Funciones base Sigmoidales:

$$\phi_j(x) = \sigma\left(\frac{x - \mu_j}{s}\right)$$

donde

$$\sigma(a) = \frac{1}{1 + \exp(-a)}.$$

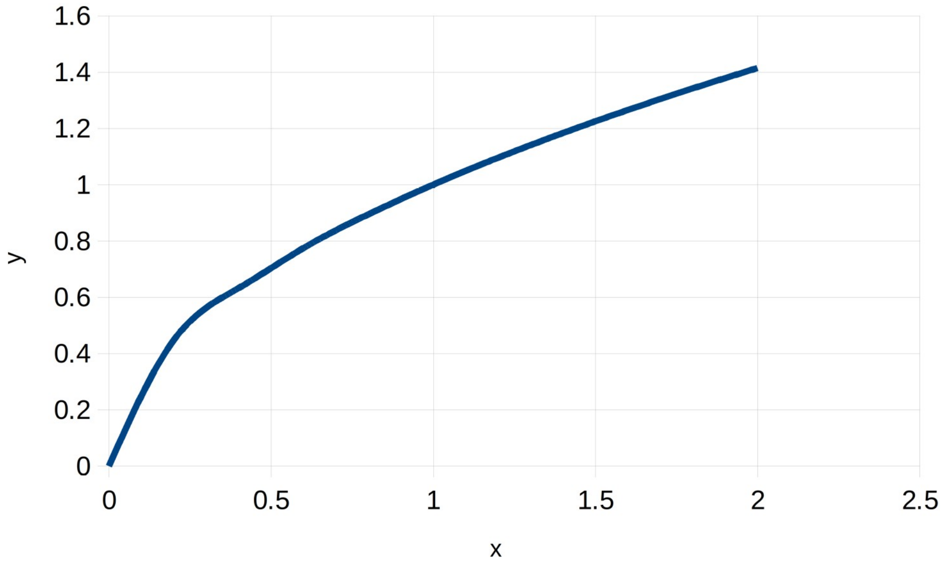


También son locales; un pequeño cambio en x solo afecta la funciones base cercanas. μ_j y s controlan la posición y la escala (pendiente).

Modelos de funciones base (5)

- Para datos que crecen hasta llegar a un cierto limite, puede ser de utilidad incluir a la función entre las funciones base

$$\varphi(x)=\sqrt{(x)}$$



Regresión exponencial

Es una adaptación de la regresión lineal para funciones exponenciales

$$f(x) = a e^{bx}$$

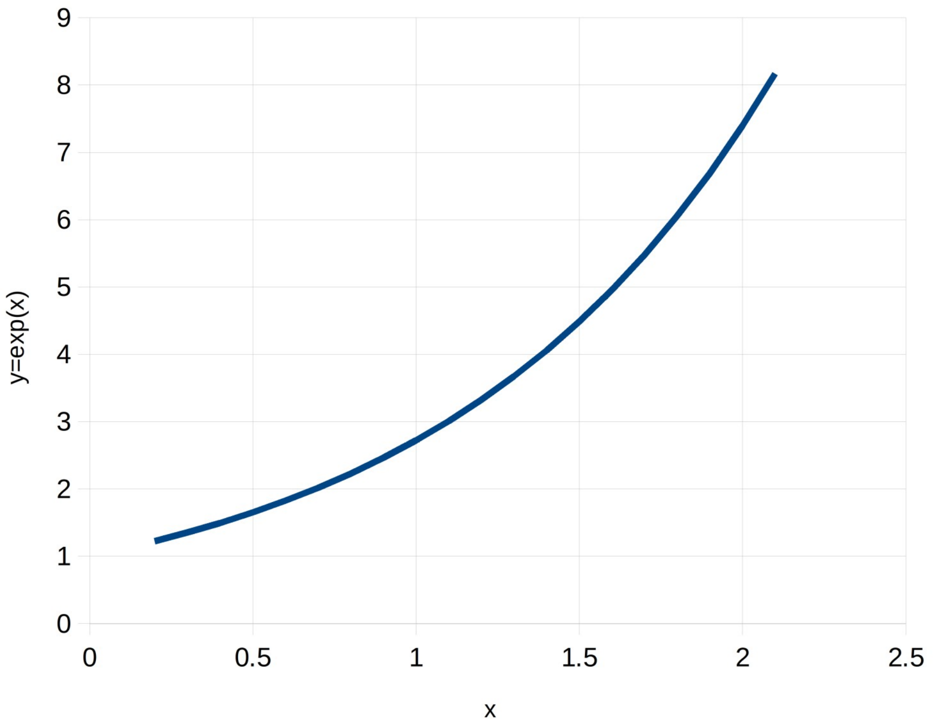
Tomado logaritmos de los targets se convierte en un problema de regresión lineal

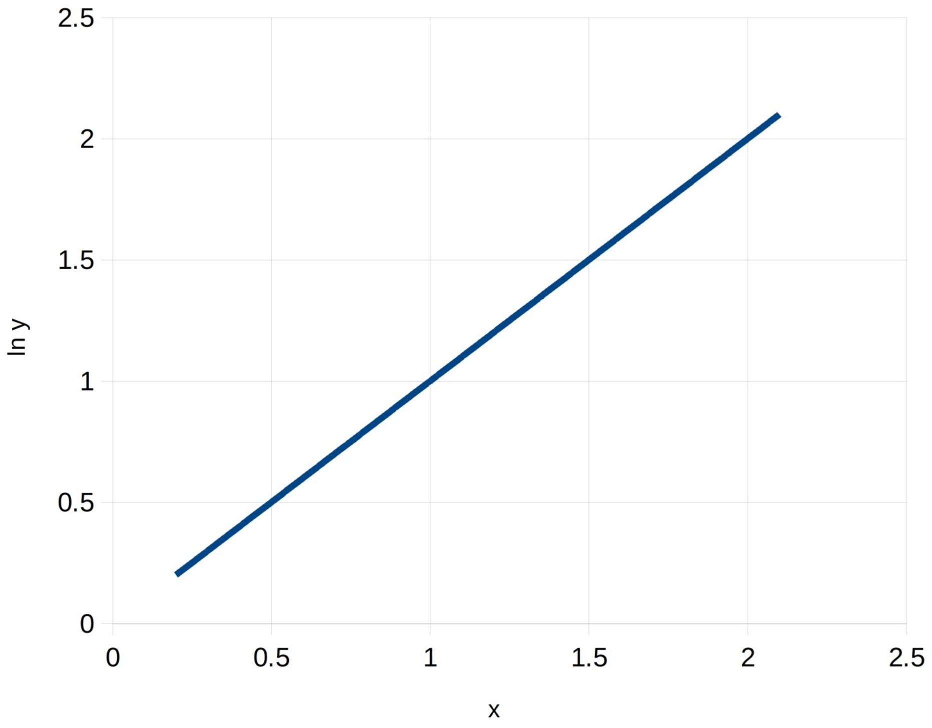
$$\ln f(x) = \ln\left(a e^{bx}\right) = \ln a + b x$$

$$y(x) = w_0 + w_1 x$$

$$a = e^{w_0} \quad b = w_1$$

Regresión exponencial (2)





Regresión de potencia

Es una adaptación de la regresión lineal para funciones de potencia

$$f(x) = a x^b$$

Tomado logaritmos de los targets y de los rasgos se convierte en un problema de regresión lineal

$$\ln f(x) = \ln\left(a x^b\right) = \ln a + b \ln x$$
$$\ln y(\ln x) = w_0 + w_1 \ln x$$
$$a = e^{w_0} \quad b = w_1$$

Regresión logarítmica

Es una adaptación de la regresión lineal para funciones de potencia

$$f(x) = a\ln(x) + b$$

Tomado logaritmos de las entradas se convierte en un problema de regresión lineal

$$f(x) = a\ln(x) + b$$
$$y(x) = w_0 + w_1 \ln x$$
$$a = w_1 \quad b = w_0$$

Múltiples rasgos

- Se deben incluir en el polinomio de grado m todas las combinaciones se productos de los rasgos para cada potencia
- Polinomio de grado 3 con 2 rasgos x_1 y x_2

$$y(\boldsymbol{x}) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1^2 + w_4 x_1 x_2 + w_5 x_2^2 + \ldots$$
$$w_6 x_1^3 + w_7 x_1^2 x_2 + w_8 x_1 x_2^2 + w_9 x_2^3$$

- Polinomio de grado 3 con D rasgos

$$y(\mathbf{x}, \mathbf{w}) = w_0 + \sum_{i=1}^{D} w_i x_i + \sum_{i=1}^{D} \sum_{j=1}^{D} w_{ij} x_i x_j + \sum_{i=1}^{D} \sum_{j=1}^{D} \sum_{k=1}^{D} w_{ijk} x_i x_j x_k$$