



Métodos Experimentales FI2003 Semestre Primavera 2009 Clase #12

Nicolás Mujica
nmujica@dfi.uchile.cl

Clase #12

- Regresión lineal: mínimos cuadrados
- Ajuste de modelos
- Ejercicios de regresión lineal
- Cifras significativas
- No solo se ajustan curvas, también superficies... (imágenes)

Regresión Lineal

- Se tiene una serie de medidas:

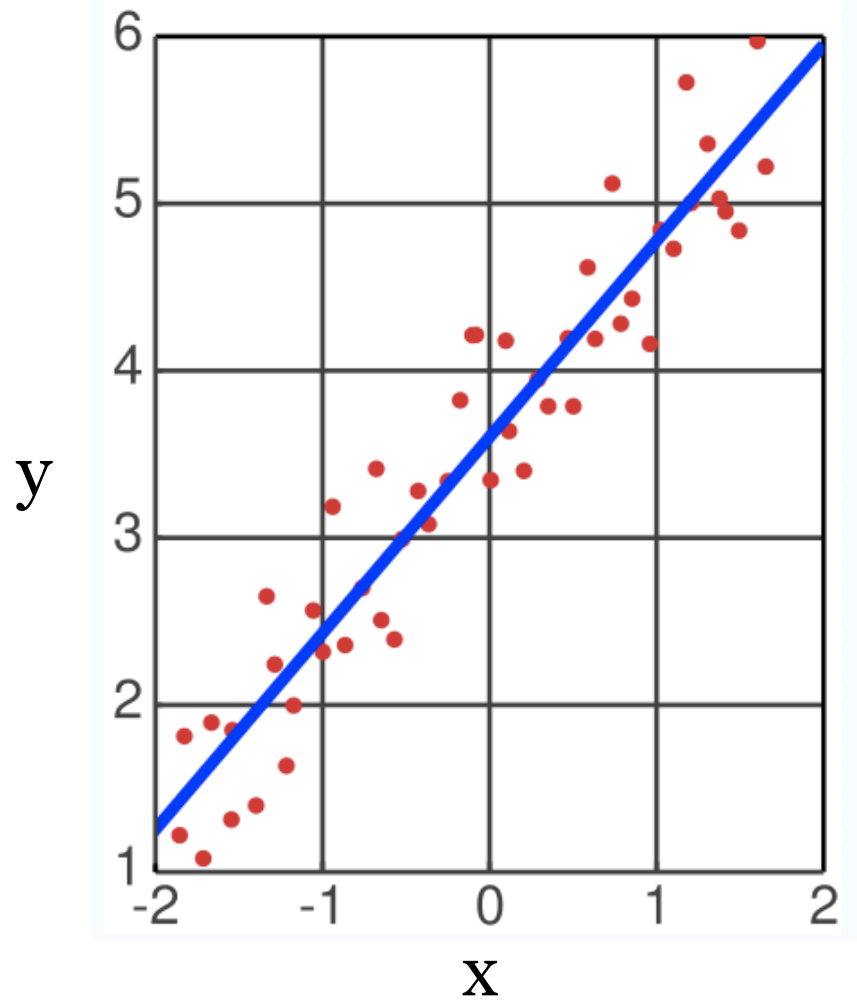
$$(x_1, y_1), (x_2, y_2), \dots (x_N, y_N)$$

Como ejemplo: x_i puede ser un tiempo i -ésimo e y_i puede ser una posición.

- Se quiere determinar si estos datos se pueden interpretar con un cierto modelo:

$$y = f(x)$$

Ajuste de modelos: caso de regresión lineal



$$y = a \cdot x + b \quad ?$$

Ajuste de modelos: caso de regresión lineal

- En el caso de un modelo lineal, el modelo es

$$y = a \cdot x + b$$

- Denotemos a los resultados de este modelo en los puntos x_i de la forma

$$y_i^m = a \cdot x_i + b$$

- Luego, lo que se busca es minimizar las “distancias”

$$|y_i^m - y_i|$$

Es decir, se buscan los parámetros a y b que permiten la “mejor” descripción de los datos medidos con el modelo considerado.

Ajuste de modelos: caso de regresión lineal

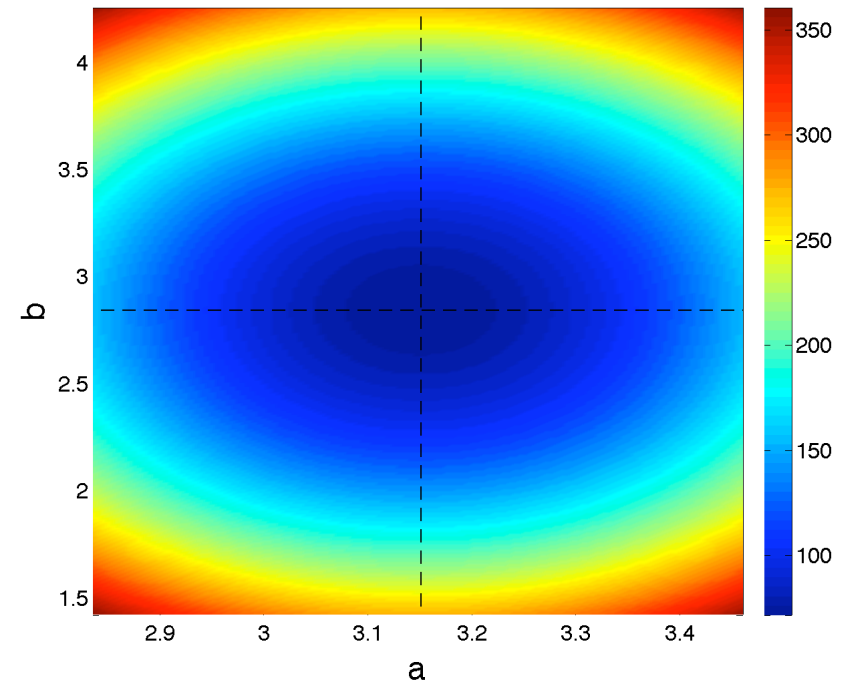
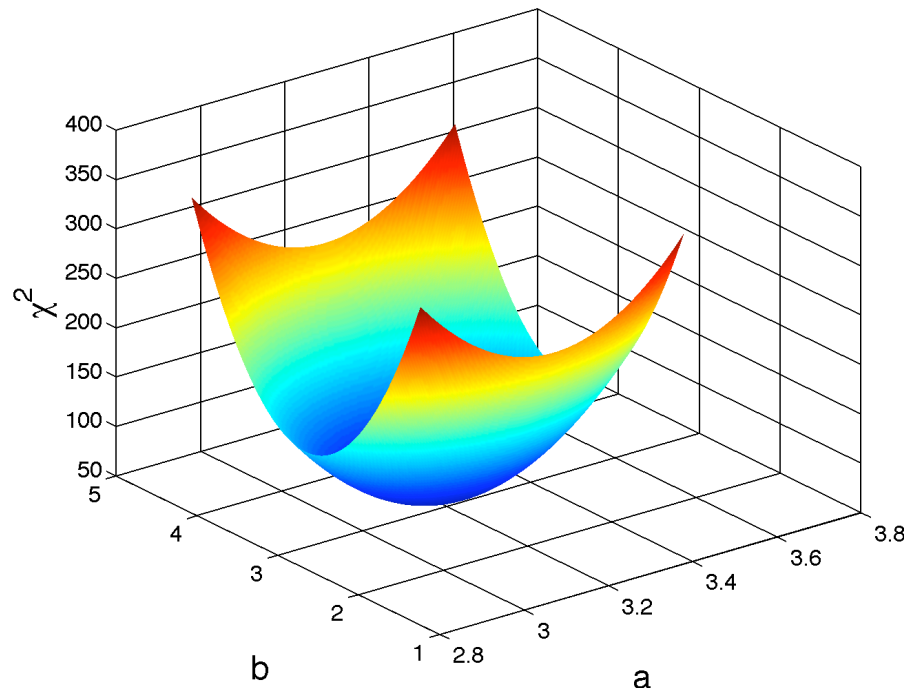
- Para ello se define:

$$\chi^2 = \sum_{i=1}^N (y_i^m - y_i)^2$$

- Y se busca minimizar esta cantidad. En el caso del modelo lineal:

$$\chi^2 = \sum_{i=1}^N (a \cdot x_i + b - y_i)^2$$

Ajuste de modelos: caso de regresión lineal



Realizar el ajuste corresponde a encontrar valores de a y b que minimizan χ^2 , es decir el mínimo global de esta superficie $\chi^2(a, b)$

Ajuste de modelos: caso de regresión lineal

- Para encontrar los mejores valores de a y b se minimiza esta cantidad, o sea imponemos:

$$\frac{\partial \chi^2}{\partial a} = \frac{\partial \chi^2}{\partial b} = 0$$

- Estas condiciones dan:

$$a \sum_{i=1}^N x_i^2 + b \sum_{i=1}^N x_i - \sum_{i=1}^N x_i \cdot y_i = 0$$

$$a \sum_{i=1}^N x_i + b \sum_{i=1}^N 1 - \sum_{i=1}^N y_i = 0$$

Ajuste de modelos: caso de regresión lineal

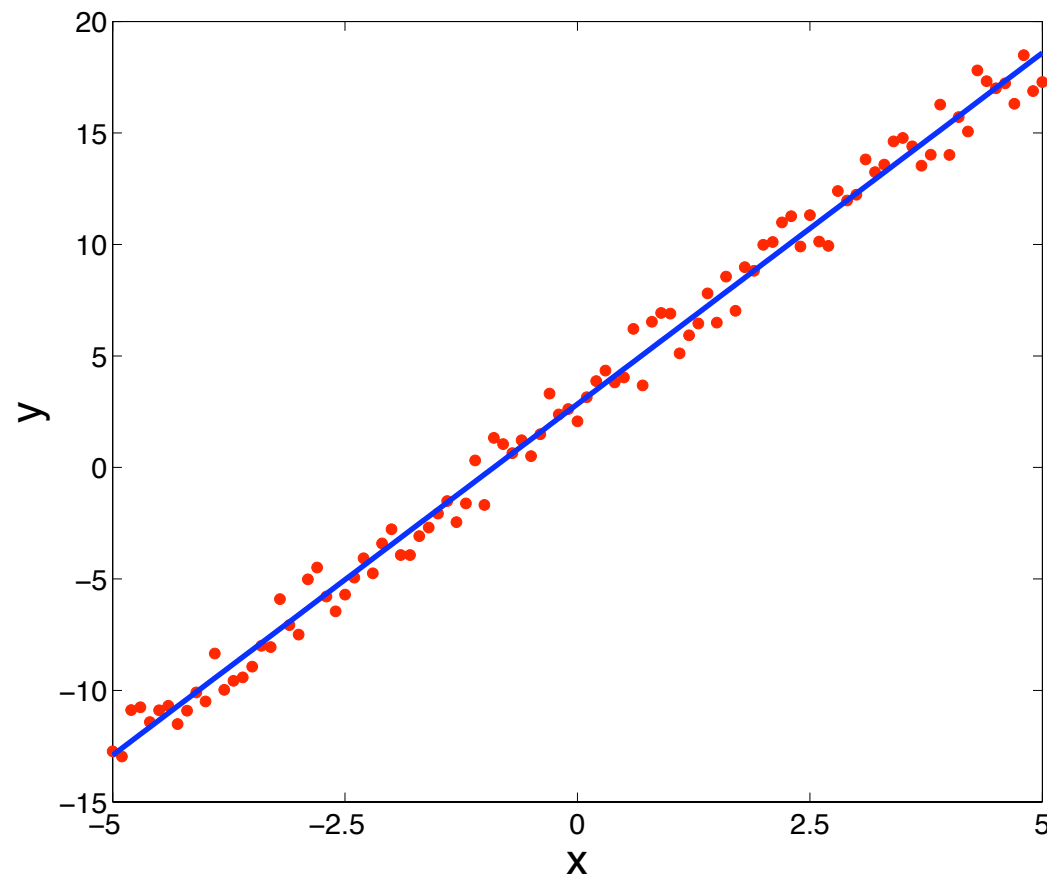
- Despejando a y b:

$$a = \frac{\left(\frac{\sum_{i=1}^N x_i \cdot y_i}{\sum_{i=1}^N x_i} - \frac{\sum_{i=1}^N y_i}{N} \right)}{\left(\frac{\sum_{i=1}^N x_i^2}{\sum_{i=1}^N x_i} - \frac{\sum_{i=1}^N x_i}{N} \right)} \quad b = \frac{\sum_{i=1}^N y_i}{N} - a \cdot \frac{\sum_{i=1}^N x_i}{N}$$

- Se define el coeficiente de regresión lineal R^2

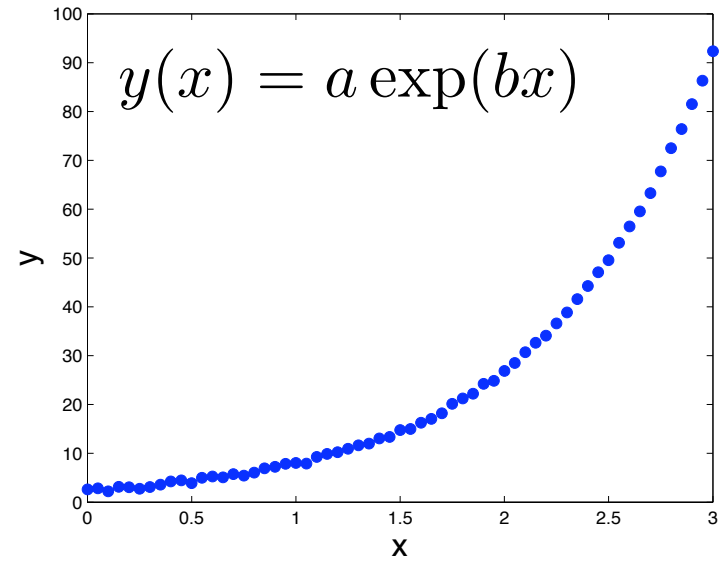
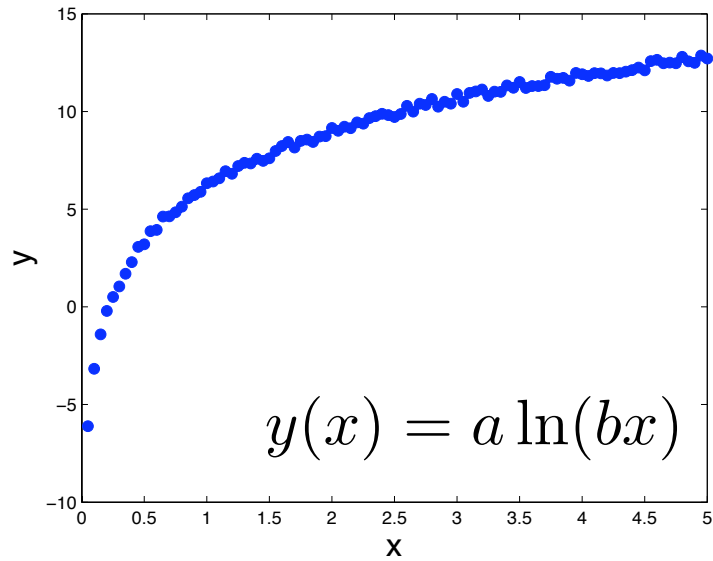
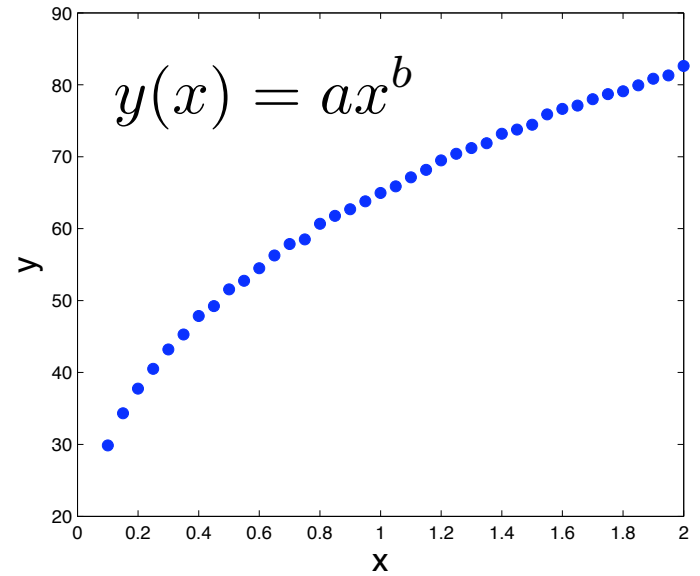
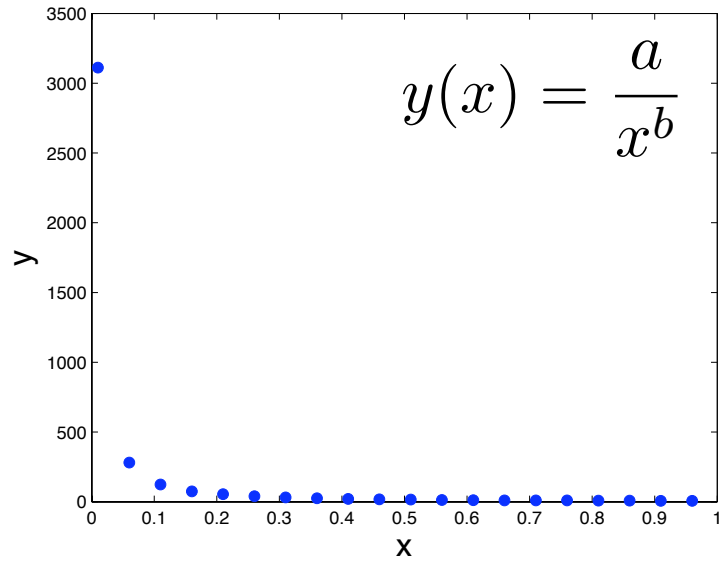
$$R^2 = 1 - \frac{\chi^2}{SST} \quad \chi^2 = \sum_{i=1}^N (y_i^m - y_i)^2 \quad \langle y_i \rangle = \sum_{i=1}^N y_i / N$$
$$SST = \sum_{i=1}^N (y_i - \langle y_i \rangle)^2$$

Ajuste de modelos: caso de regresión lineal



$$a = 3.15, \quad b = 2.84, \quad R^2 = 0.992$$

Se puede hacer mas?



Ejemplos

Parte A - Regresión Lineal: (2 pts) En un experimento se mide la amplitud presión acústica P en un tubo circular en función de la posición x a lo largo del tubo. Los datos se presentan en la tabla 1.

P (Pa)	8.35	4.37	3.29	1.61	0.58
x (cm)	3.0	13.5	17.4	30.0	46.2

Tabla 1

P_o	α	R^2

Tabla 2

Sabiendo que presión y la distancia están relacionadas por la ecuación $P = P_o \cdot e^{\alpha x}$, determine los valores de P_o , α , y el coeficiente de regresión R^2 asociado. Complete la tabla 2 y no olvide unidades. El desarrollo de esta parte debe ser entregado en una hoja aparte.



Ejemplos

FI2003 - Métodos Experimentales - secciones 3 y 4
Ejercicio N° 4 - optativo
Departamento de Física
Facultad de Ciencias Físicas y Matemáticas
Universidad de Chile

Prof. Nicolás Mujica
23 de junio de 2009
Tiempo: 45 minutos

P1. Se tiene un canal largo y angosto lleno parcialmente con agua el cual es vibrado verticalmente de manera sinusoidal con una pequeña amplitud. Bajo ciertas condiciones se observa una estructura localizada llamada *Soliton hidrodinámico no propagativo*, como se muestra en la figura 1, la cual oscila entre la pared frontal y trasera de canal.

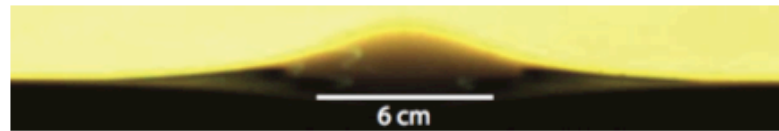


Figura 1: Fotografía de un solitón hidrodinámico no propagativo.

En ocasiones se forman dos de estos solitones que interactúan de manera repulsiva, de modo que la distancia que los separa aumenta en función del tiempo. Una medición de la distancia Δ entre los máximos de dos de estos solitones en función del tiempo t es reportada en el cuadro 1.

Cuadro 1: Distancia entre máximos de los solitones en función del tiempo.

t (s)	8.3	13.2	25.5	43.0	55.1	67.8
Δ (cm)	5.2	6.1	7.5	8.5	9.0	9.4

Ejemplos

- (a) Haga un gráfico de Δ versus t rotulando debidamente los ejes. (1 pt)
- (b) Se consideran tres opciones para describir la evolución de Δ versus t :
- (i) $\Delta = a \cdot t + b$,
 - (ii) $\ln(\Delta) = a \cdot t + b$,
 - (iii) $\Delta = a \cdot \ln(t) + b$,

Sabiendo que $a > 0$, pues los solitones se alejan en función del tiempo, ¿Cuál de las opciones anteriores describe mejor el comportamiento observado? Justifique su respuesta en base a lo graficado en la parte (a). (2 pts)

- (c) Considerando el modelo que mejor describe el sistema, determine los valores de a , b , y el coeficiente de regresión R^2 asociado del ajuste lineal correspondiente. (3 pts)

Nota: si no sabe justificar la parte (b) puede hacer la parte (c) para cada opción de evolución temporal y justificar su respuesta en función del R^2 obtenido.

Cifras significativas

Se realizan una serie de medidas de una cantidad física

1,23245

1,23542

1,23910

1,23145

1,23132

1,22981

1,23437

1,24014

Cifras significativas

Se realizan una serie de medidas de una cantidad física

1,23245

1,23542

1,23910

1,23145

1,23132

1,22981

1,23437

1,24014

Matlab nos da:

Promedio = 1,2342575

Desviación estándar = 0,00376024220345

Tiene sentido tantas cifras?

Cifras significativas

Se realizan una serie de medidas de una cantidad física

1,23245

1,23542

1,23910

1,23145

1,23132

1,22981

1,23437

1,24014

Matlab nos da:

Promedio = 1,2342575

Desviación estándar = 0,00376024220345

Tiene sentido tantas cifras?

Cifras significativas

Se realizan una serie de medidas de una cantidad física

1,23245

1,23542

1,23910

1,23145

1,23132

1,22981

1,23437

1,24014

Matlab nos da:

Promedio = 1,2342575

Desviación estándar = 0,00376024220345

Tiene sentido tantas cifras?

Cifras significativas

Se realizan una serie de medidas de una cantidad física

1,23245

1,23542

1,23910

1,23145

1,23132

1,22981

1,23437

1,24014

Matlab nos da:

Promedio = 1,234

Desviación estándar = 0,004

Decimos entonces que la cantidad
tiene 4 cifras significativas

Cifras significativas: más ejemplos

$$T = 24.12 \pm 0.21 \text{ }^{\circ}\text{C} \rightarrow T = 24.1 \pm 0.2 \text{ }^{\circ}\text{C}$$

$$m = 0.019812 \pm 0.000267 \text{ kg} \rightarrow m = 0.0198 \pm 0.0003 \text{ kg}$$

$$I = 12.4 \pm 1.23 \text{ A} \rightarrow I = 12 \pm 1 \text{ A}$$