



LOYOLA
Másteres

TEMA 2

Introducción Práctica a la Regresión Lineal

2022/2023

ASIGNATURA Modelos de Regresión

¿Cuánto vale mi casa?



**Quiero poner
mi casa a la
venta**

¿Cuánto vale mi casa?



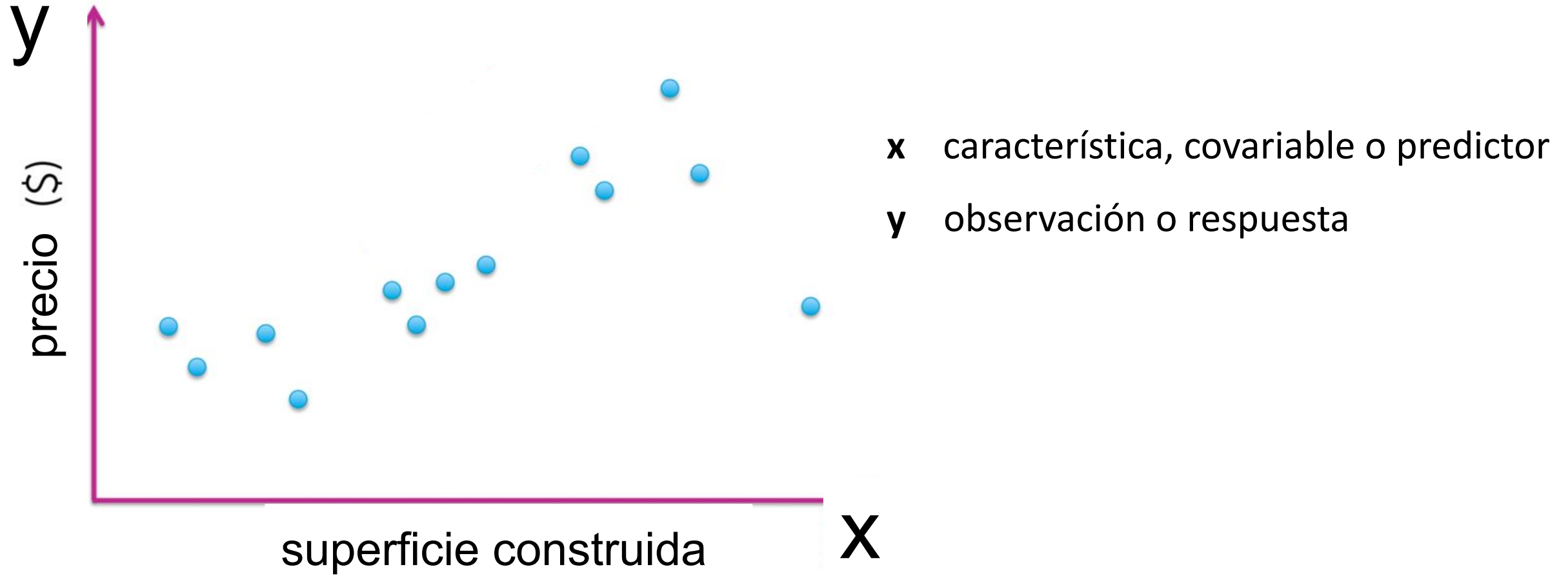
Busco ventas recientes en mi barrio

¿Por cuánto se vendieron otras viviendas?

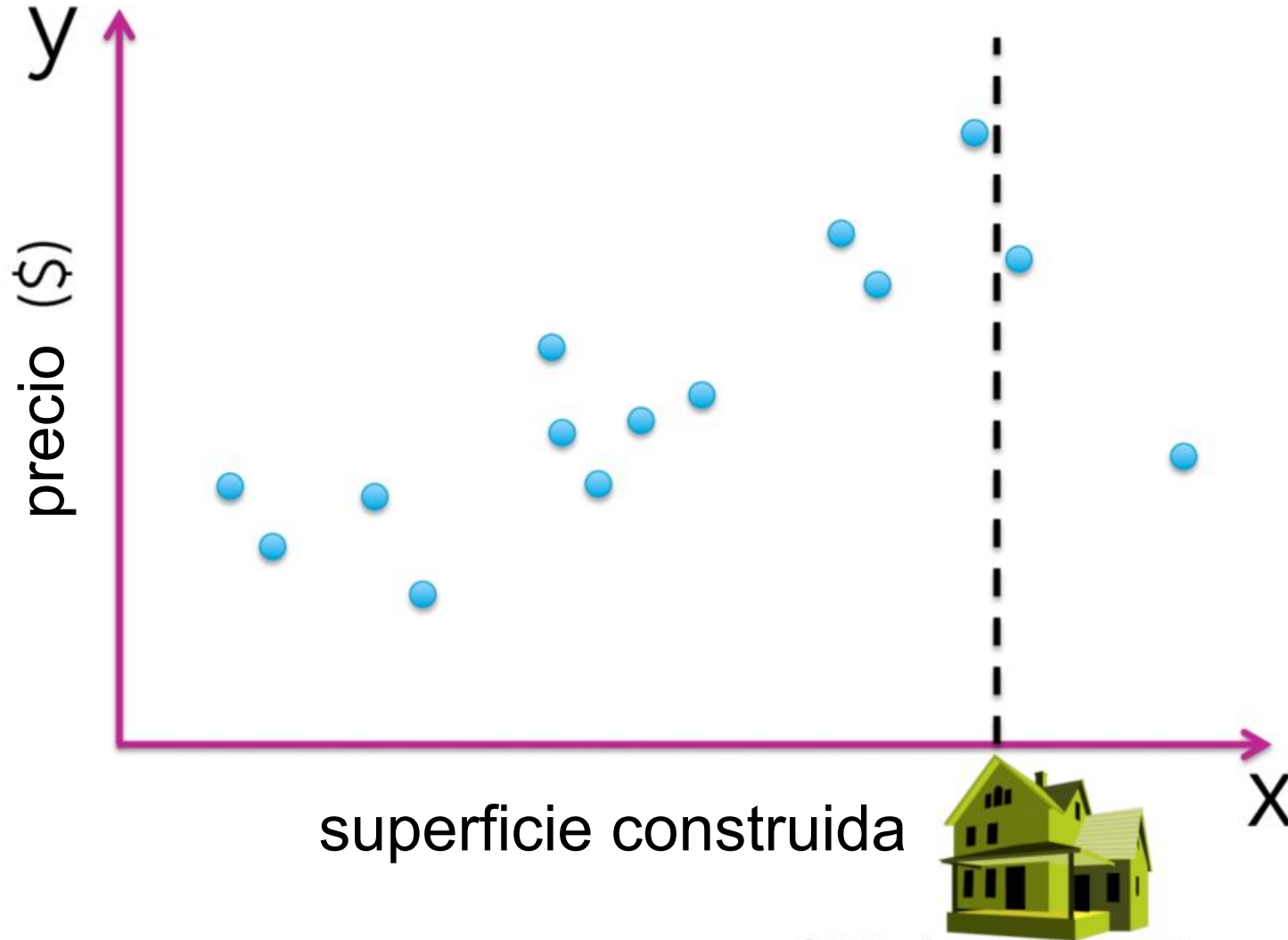


Mostremos las ventas más recientes

Últimos 2 años

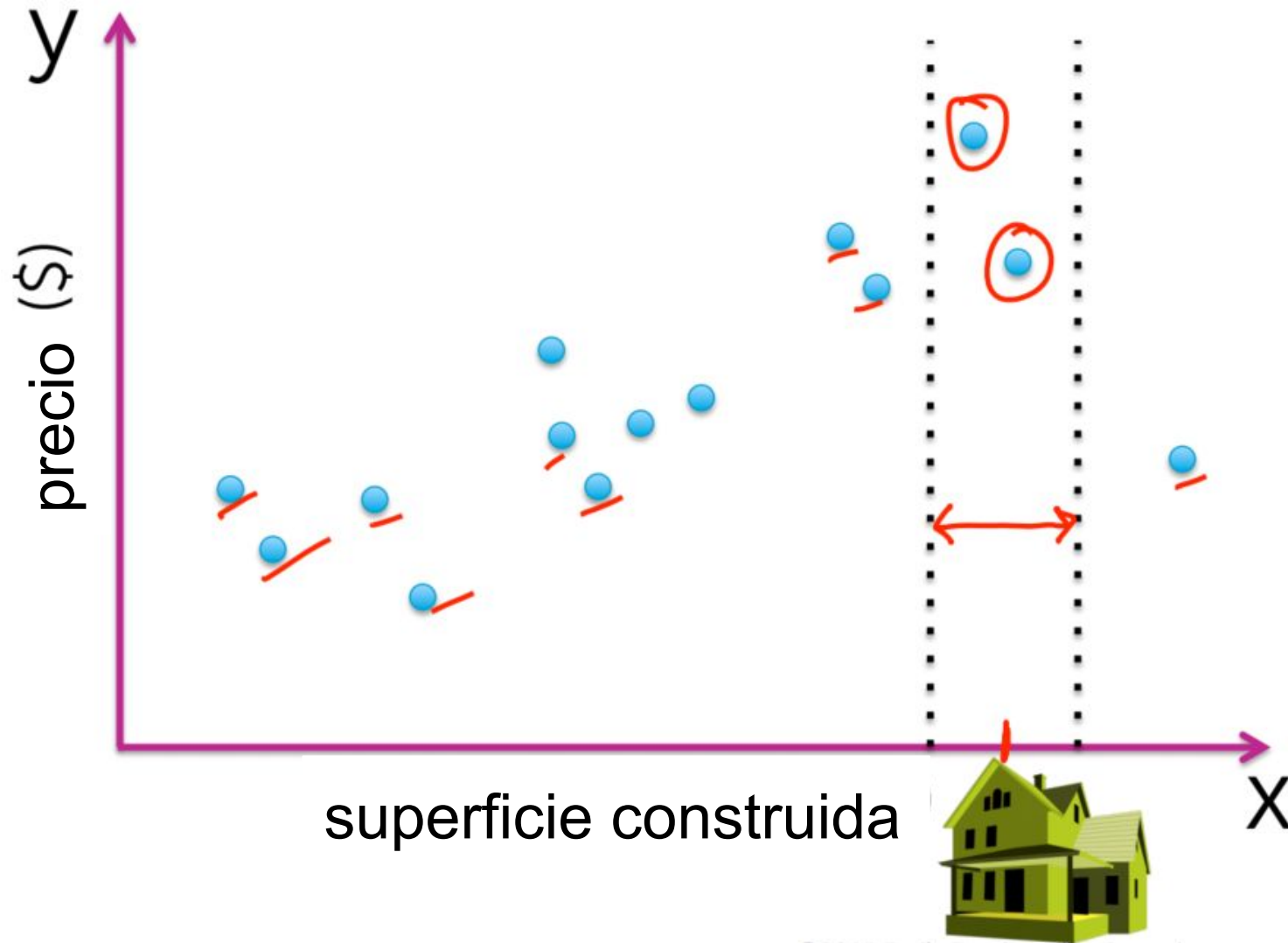


Predicción basada en casas similares



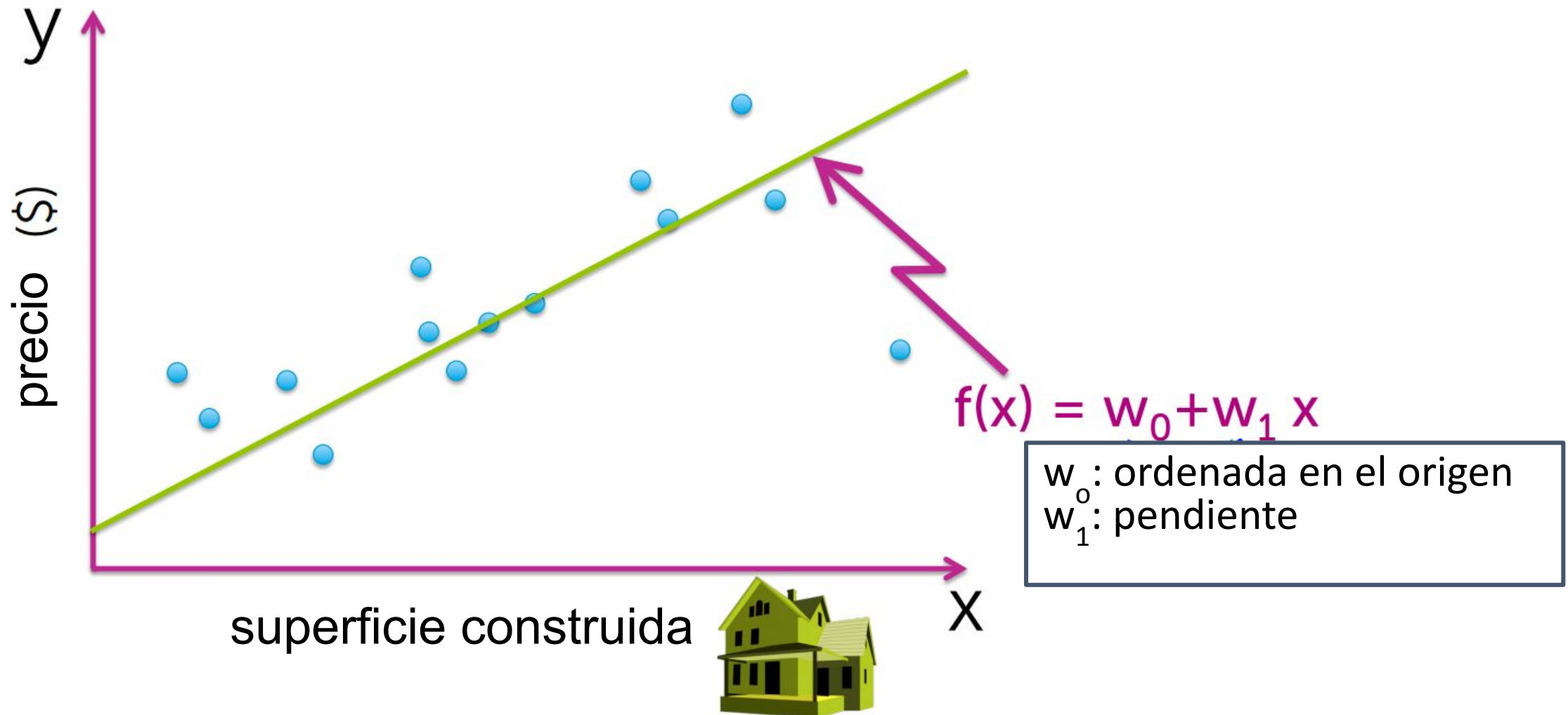
Ninguna de las casas vendidas tenía exactamente la misma superficie

Predicción basada en casas similares



- Estima con la media en un rango cercano
- **¡Sólo 2 casas!**
- Tenemos que usar la info de todas las ventas

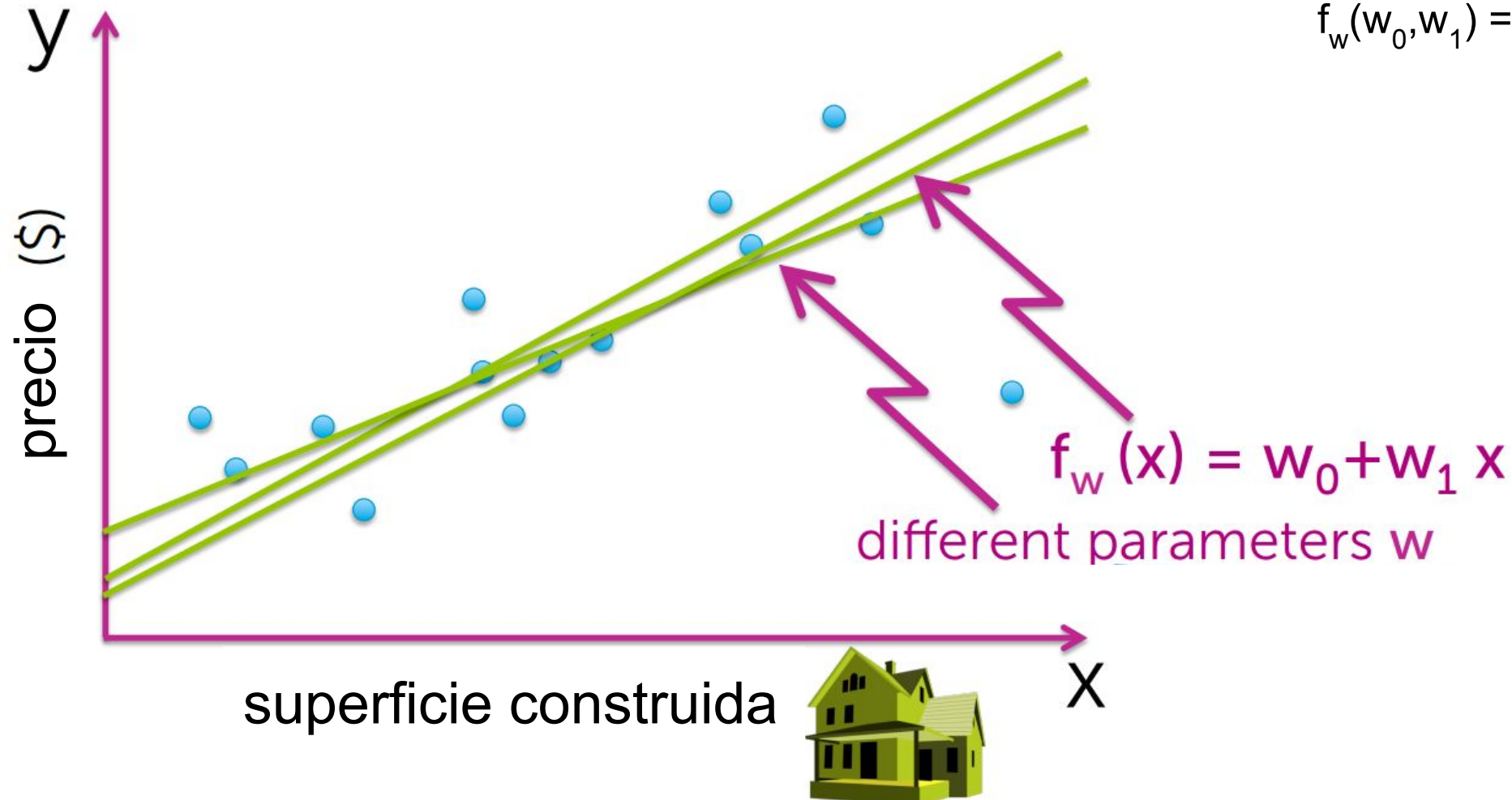
Modelo de regresión lineal



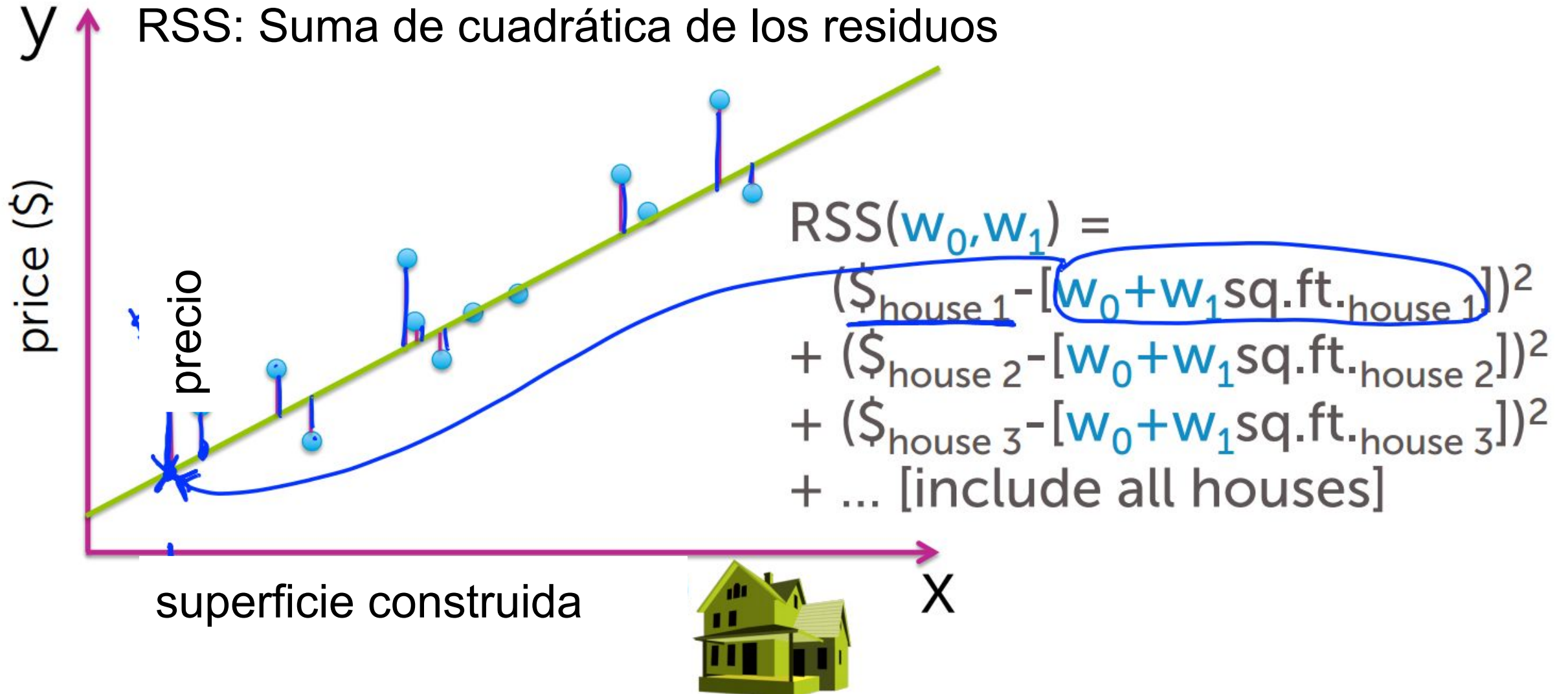
Ajuste a la mejor recta

Función parametrizada para optimizar su ajuste a la nube de puntos:

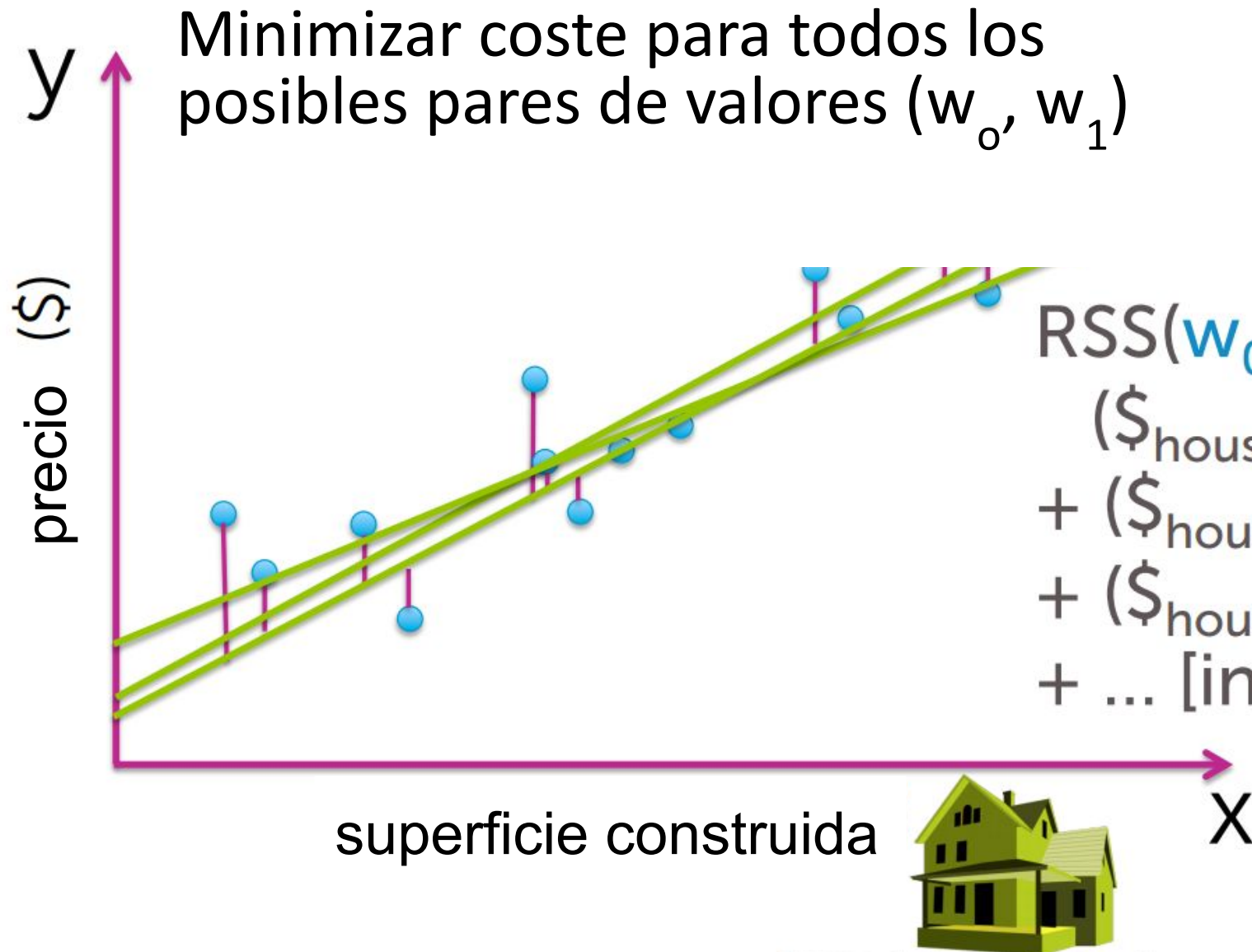
$$f_w(w_0, w_1) = w_0 + w_1 x$$



Coste de una instancia de la función



Búsqueda de la “mejor” línea



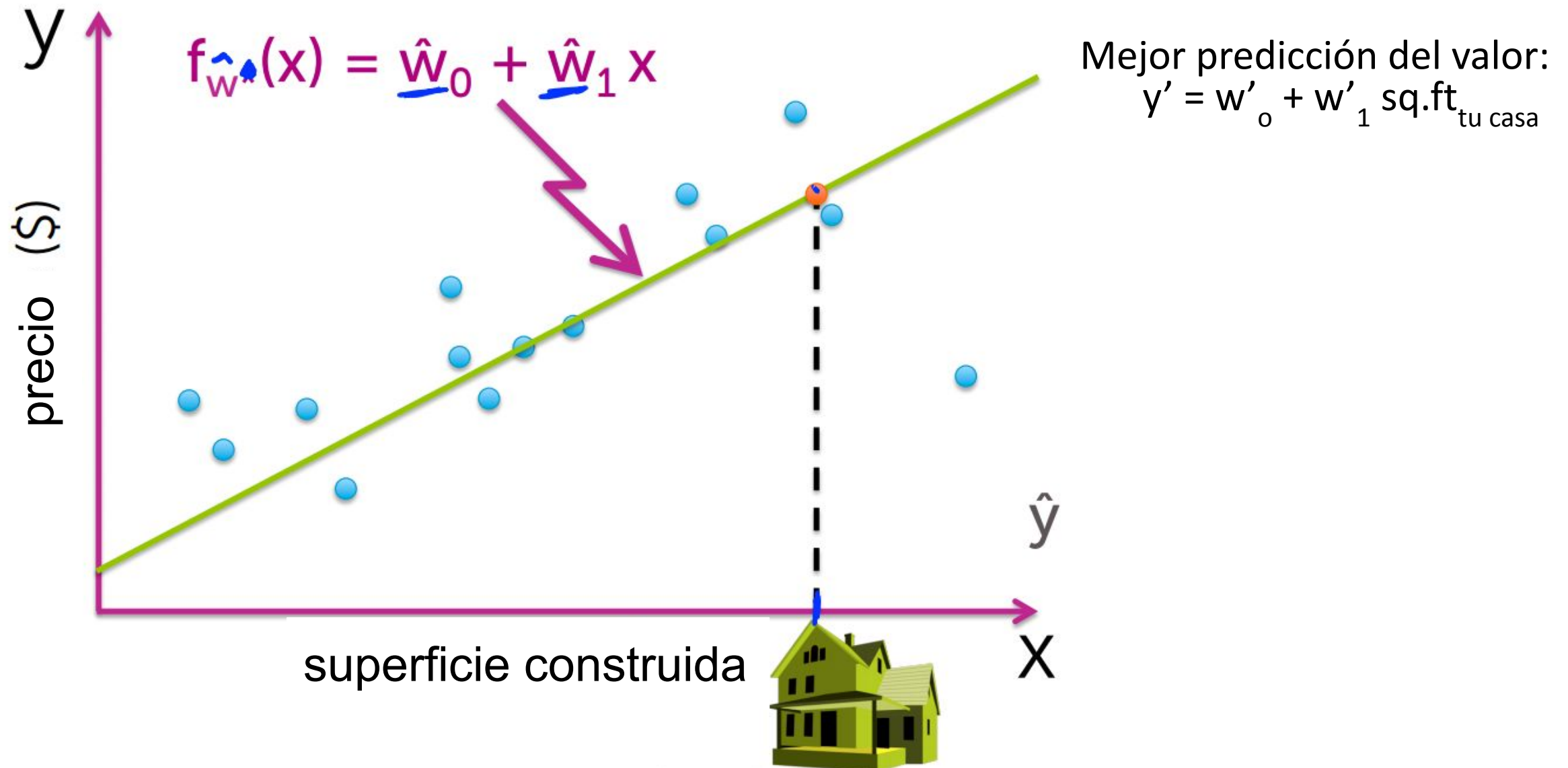
$$RSS(w_0, w_1) =$$

$$\begin{aligned} & (\$_{\text{house 1}} - [w_0 + w_1 \text{sq.ft.}_{\text{house 1}}])^2 \\ & + (\$_{\text{house 2}} - [w_0 + w_1 \text{sq.ft.}_{\text{house 2}}])^2 \\ & + (\$_{\text{house 3}} - [w_0 + w_1 \text{sq.ft.}_{\text{house 3}}])^2 \\ & + \dots [\text{include all houses}] \end{aligned}$$

↓

$$\hat{W} = (\hat{w}_0, \hat{w}_1)$$

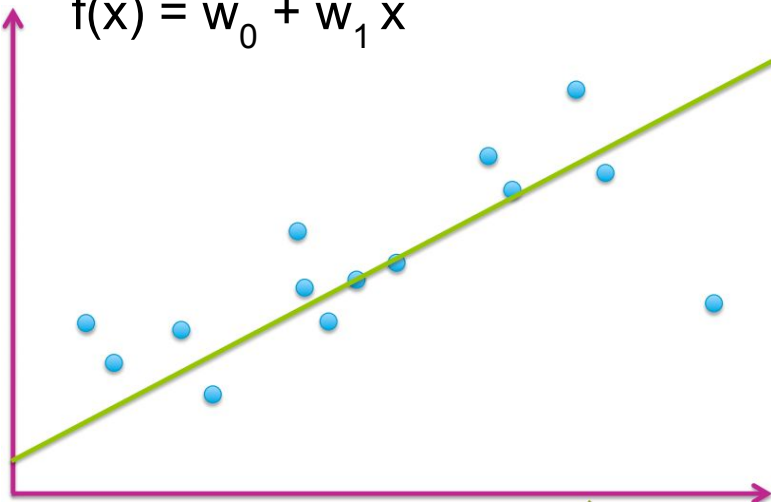
Predicción del precio de tu vivienda



Mejor curva de ajuste

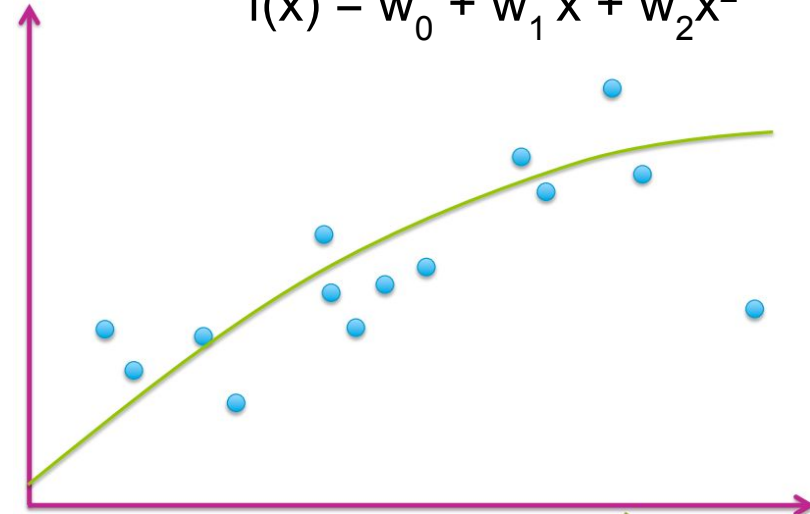
relación lineal

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

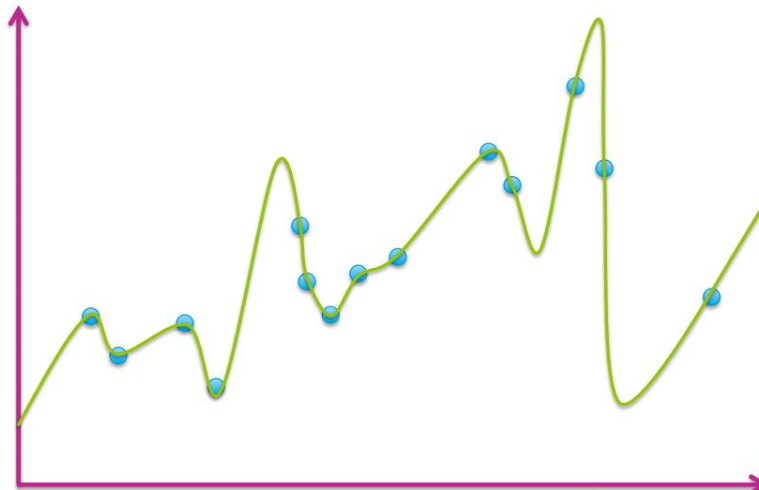


relación cuadrática

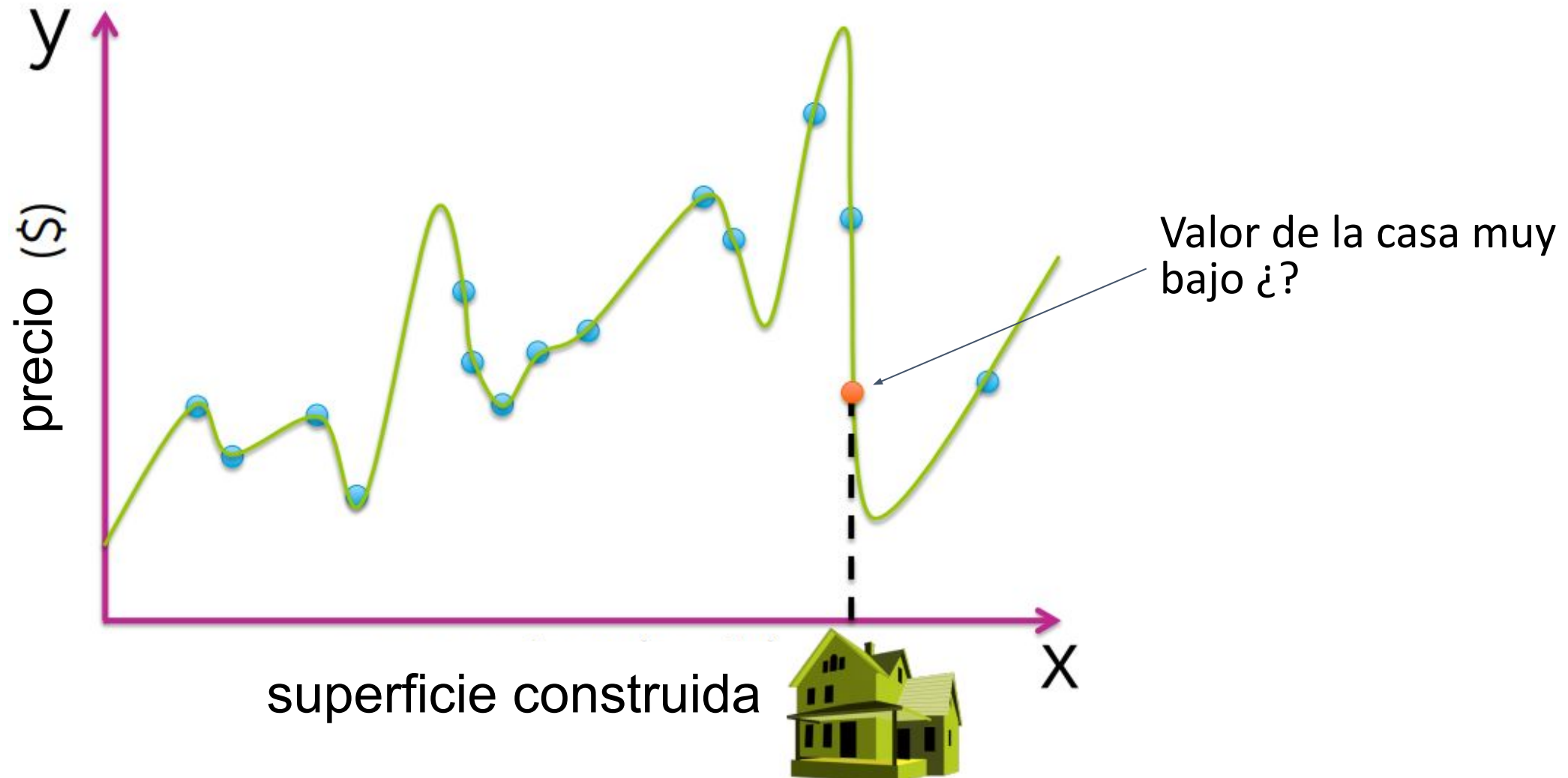
$$f(x) = w_0 + w_1 x + w_2 x^2$$



polinomio de grado 'n'



¿Es creíble esta predicción?



Partición training / test sets

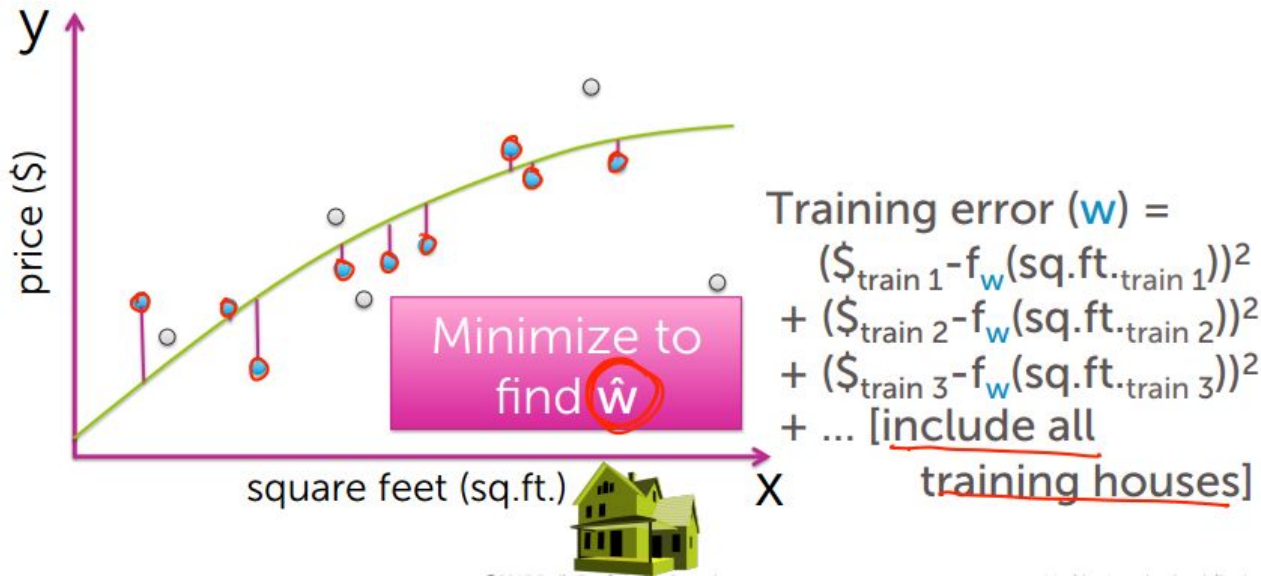


- Training set 80%
- Test set 20%

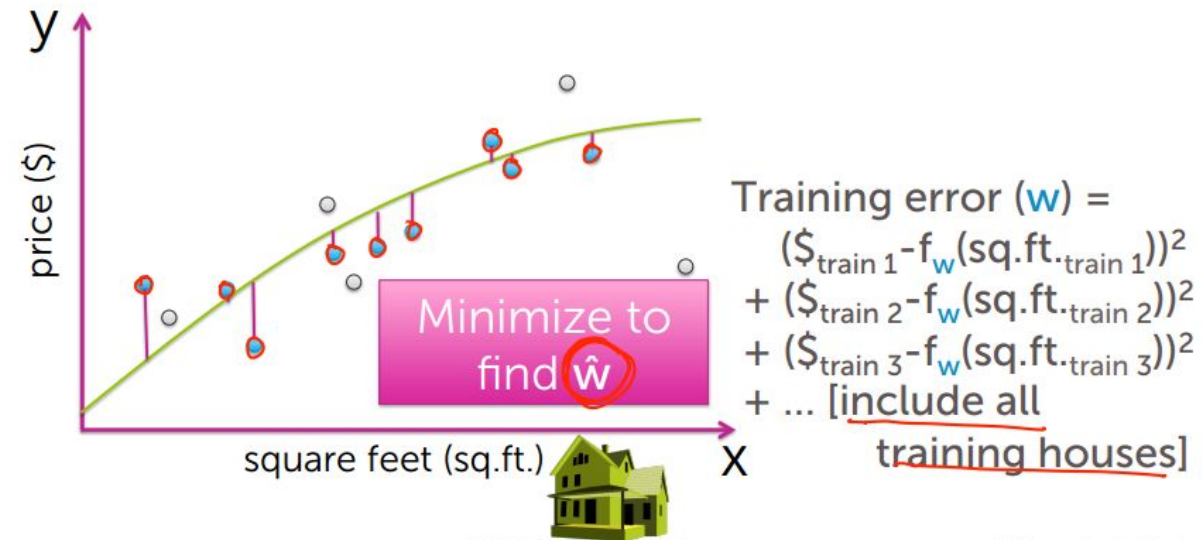
Terminology: – training set
– test set



Training error



Test error



Overfitting

