

Affine functions

- ▶ a function that is linear plus a constant is called *affine*
- ▶ general form is $f(x) = a^T x + b$, with a an n -vector and b a scalar
- ▶ a function $f : \mathbf{R}^n \rightarrow \mathbf{R}$ is affine if and only if

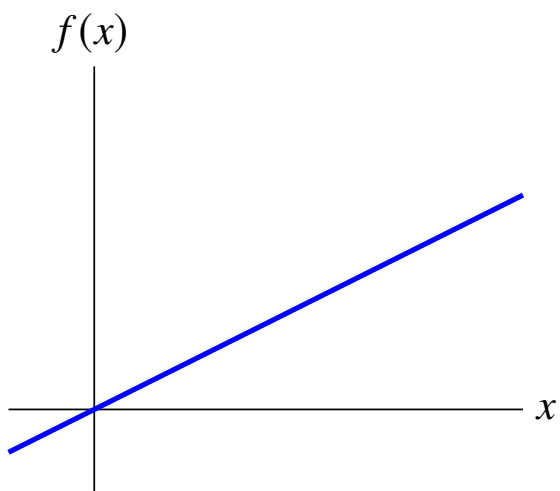
$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

holds for all α, β with $\alpha + \beta = 1$, and all n -vectors x, y

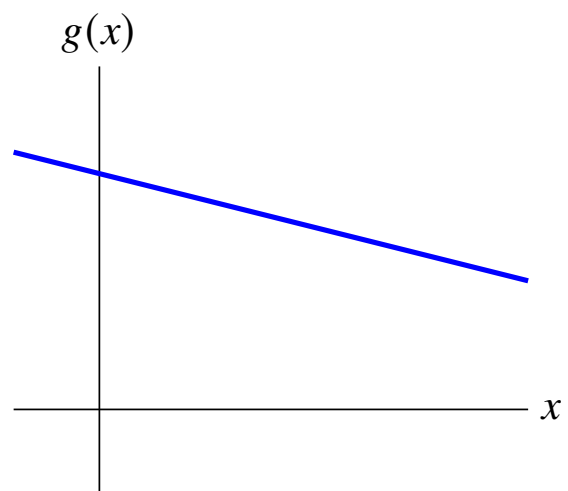
- ▶ sometimes (ignorant) people refer to affine functions as linear

Linear versus affine functions

f is linear



g is affine, not linear



Outline

Linear and affine functions

Taylor approximation

Regression model

Regression model

- ▶ *regression model* is (the affine function of x)

$$\hat{y} = x^T \beta + v$$

- ▶ x is a feature vector; its elements x_i are called *regressors*
- ▶ n -vector β is the *weight vector*
- ▶ scalar v is the *offset*
- ▶ scalar $\hat{y}$ is the *prediction*
(of some actual outcome or *dependent variable*, denoted y)

Example

- ▶ y is selling price of house in \$1000 (in some location, over some period)

- ▶ regressor is

$$x = (\text{house area, \# bedrooms})$$

(house area in 1000 sq.ft.)

- ▶ regression model weight vector and offset are

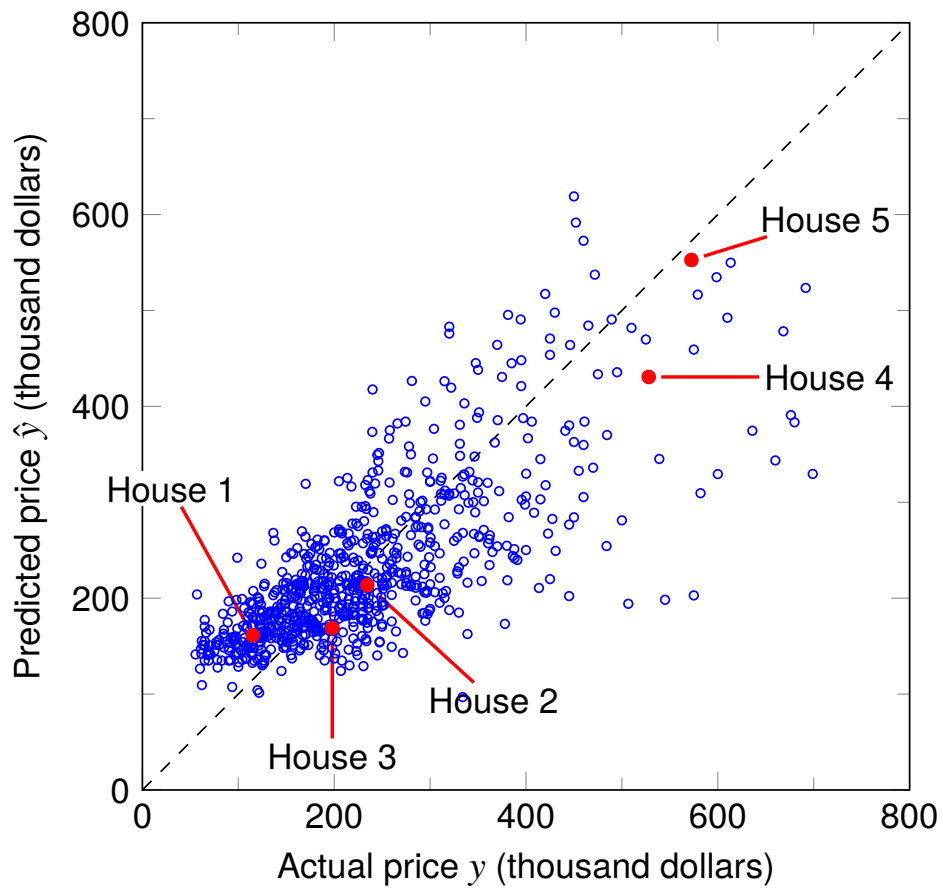
$$\beta = (148.73, -18.85), \quad \nu = 54.40$$

- ▶ we'll see later how to guess β and ν from sales data

Example

House	x_1 (area)	x_2 (beds)	y (price)	$\hat{y}$ (prediction)
1	0.846	1	115.00	161.37
2	1.324	2	234.50	213.61
3	1.150	3	198.00	168.88
4	3.037	4	528.00	430.67
5	3.984	5	572.50	552.66

Example



3. Norm and distance

Outline

Norm

Distance

Standard deviation

Angle

Norm

- ▶ the *Euclidean norm* (or just *norm*) of an n -vector x is

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{x^T x}$$

- ▶ used to measure the size of a vector
- ▶ reduces to absolute value for $n = 1$

Properties

for any n -vectors x and y , and any scalar β

- ▶ *homogeneity*: $\|\beta x\| = |\beta| \|x\|$
- ▶ *triangle inequality*: $\|x + y\| \leq \|x\| + \|y\|$
- ▶ *nonnegativity*: $\|x\| \geq 0$
- ▶ *definiteness*: $\|x\| = 0$ only if $x = 0$

easy to show except triangle inequality, which we show later

RMS value

- ▶ *mean-square value* of n -vector x is

$$\frac{x_1^2 + \cdots + x_n^2}{n} = \frac{\|x\|^2}{n}$$

- ▶ *root-mean-square value* (RMS value) is

$$\mathbf{rms}(x) = \sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}} = \frac{\|x\|}{\sqrt{n}}$$

- ▶ $\mathbf{rms}(x)$ gives ‘typical’ value of $|x_i|$
- ▶ e.g., $\mathbf{rms}(\mathbf{1}) = 1$ (independent of n)
- ▶ RMS value useful for comparing sizes of vectors of different lengths