Scalar-valued Functions (Linear and Affine)

CE282: Linear Algebra

Computer Engineering Department

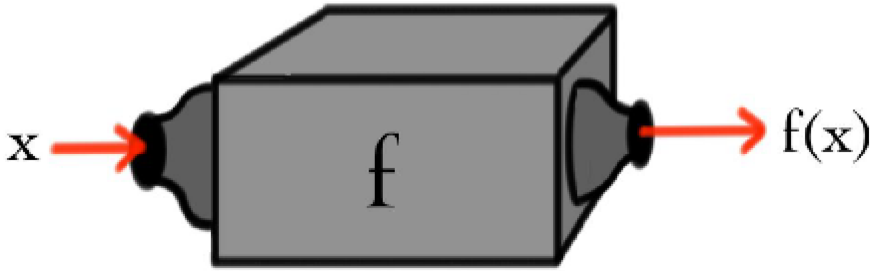
Sharif University of Technology

Hamid R. Rabiee

Maryam Ramezani

What are Functions?

- ❑ Think of a function as a machine f into which one may feed a real number. For each input x this machine outputs a $f(x)$.



(A) What number x satisfies $10x = 3$?

(B) What 3-vector v satisfies $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \times v = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$?

(C) What polynomial p satisfies $\int_{-1}^{1} p(y)\, dy = 0$ and $\int_{-1}^{1} yp(y)\, dy = 1$?

(D) What power series $f(x)$ satisfies $x\frac{d}{dx}f(x) - 2f(x) = 0$?

(E) What number x satisfies $4x^2 = 1$?

What vector X satisfies $f(\mathrm{X}) = \mathrm{B}$?

u_1	v_1	
u_2	v_2	
u_3	v_3	$u_2v_3 - u_3v_2$
u_1	v_1	$u_3v_1 - u_1v_3$
u_2	v_2	$u_1v_2 - u_2v_1$

What are Functions?

Note

- Linear and affine functions in this session are scalar-valued. We focus on the linear function machine of the previous slide, which outputs are scalar values. Remains will discuss later.

What are Linear Functions?

- ❑ $f: R^n \to R$ means that f is a function that maps real n-vectors to real numbers
- ❑ $f(x)$ is the value of function f at x (x is referred to as the argument of the function).
- ❑ $f(x) = (x_1, x_2, \dots, x_n)$: argument

Definition

A function $f: R^n \to R$ is linear if it satisfies the following two properties:

- ❑ Additivity: For any n-vector x and y, $f(x + y) = f(x) + f(y)$
- ❑ Homogeneity: For any n-vector x and any scalar $\alpha \in R$: $f(\alpha x) = \alpha f(x)$

Superposition property:

Definition

Superposition property:

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

Note

- A function that satisfies the superposition property is called linear

Definition

- **Additivity:**

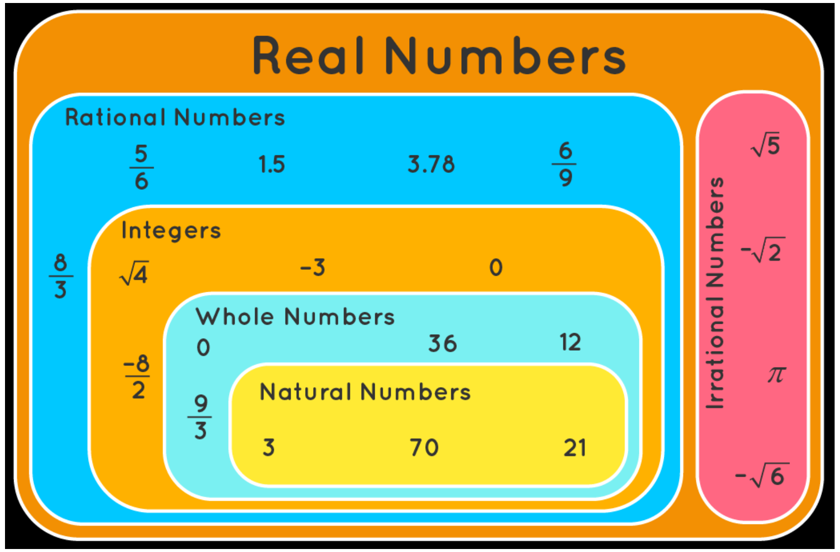
For any n-vector x and y, $f(x + y) = f(x) + f(y)$

- **Homogeneity:**

For any n-vector x and any scalar $\alpha \in R$: $f(\alpha x) = \alpha f(x)$

Counterexample:

$f(x) = f\left(a + \sqrt{5}\, b\right) \to a + b + \sqrt{5}\, b$



What are Linear Functions?

- If a function f is linear, superposition extends to linear combinations of any number of vectors:

$$f(\alpha_1 x_1 + \cdots + \alpha_k x_k) = \alpha_1 f(x_1) + \cdots + \alpha_k f(x_k)$$

Inner product is Linear Function?

Theorem

A function defined as the inner product of its argument with some fixed vector is linear.

- Proof?

$$f(x) = a^T x = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

What are Linear Functions?

Theorem

If a function is linear, then it can be expressed as the inner product of its argument with some fixed vector.

- Proof?

What are Linear Functions?

Theorem

The representation of a linear function f as $f(x) = a^T x$ is unique, which means that there is only one vector a for which $f(x) = a^T x$ holds for all x.

Affine Function

Definition

A function $f: R^n \rightarrow R$ is affine if and only if it can expressed as $f(x) = a^T x + b$ (linear function plus a constant (**offset**))

- Superposition property for affine function which is called **restricted superposition**

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)\,,\ \alpha + \beta = 1$$

Affine Function

Theorem

Any scalar-valued function that satisfies the restricted superposition property is affine.

Conclusion

Every affine function can be written as $f(x) = a^T x + b$ with:

$$a^T = [f(e_1) - f(0), f(e_2) - f(0), \dots, f(e_n) - f(0)]$$
$$b = f(0)$$

Conclusion

We can write linear and affine functions in two methods:

- Method 1:
  - Linear:

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n), \forall \alpha_1, \dots, \alpha_n$$

  - Affine:

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) , \alpha_1 + \cdots + \alpha_n = 1$$

- Method 2:
  - Linear:

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

  - Affine:

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

Definition

In many applications, scalar-valued functions of n variables, or relations between n variables and a scalar one, can be approximated as linear or affine functions, which is called "**Model**".

Scalar-valued function of a scalar

- Derivative of function $f: R \to R$ at the point $(z, f(z))$:

$$\lim_{t \to 0} \frac{f(z+t) - f(z)}{t}$$

  - It gives the slope of the graph of f at the point $(z, f(z))$.
  - $f'(z)$ is a scalar-valued function of a scalar variable

- The partial derivative of function $f: R^n \to R$ at the point z, with respect to its ith argument

$$\frac{\partial f}{\partial x_i}(z) = \lim_{t \to 0} \frac{f(z_1, \dots, z_{i-1}, z_i + t, z_{i+1}, \dots, z_n) - f(z)}{t} = \lim_{t \to 0} \frac{f(z + te_i) - f(z)}{t}$$

  - The partial derivative is the derivative with respect to the i −th argument, with all other arguments fixed.

- **Gradient**: The partial derivatives of $f(x)$ with respect to its n arguments can be collected into an n vector called the gradient of $f(x)$ (at point z):

$$\nabla f(z) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(z) \\ \vdots \\ \frac{\partial f}{\partial x_n}(z) \end{bmatrix}$$

Theorem

Gradient of a combination of functions:

$$f(z) = ag(z) + bh(z)$$

$$\nabla f(z) = a\nabla g(z) + b\nabla h(z)$$

How to find an approximate affine model

- $f: R^n \rightarrow R$ is differentiable: its partial derivatives exist

Definition

The (first-order) Taylor approximation of f near (or at) the point z:

$$\hat{f}(x) = f(z) + \frac{\partial f}{\partial x_1}(x_1 - z) + \cdots + \frac{\partial f}{\partial x_n}(x_n - z)$$

Example

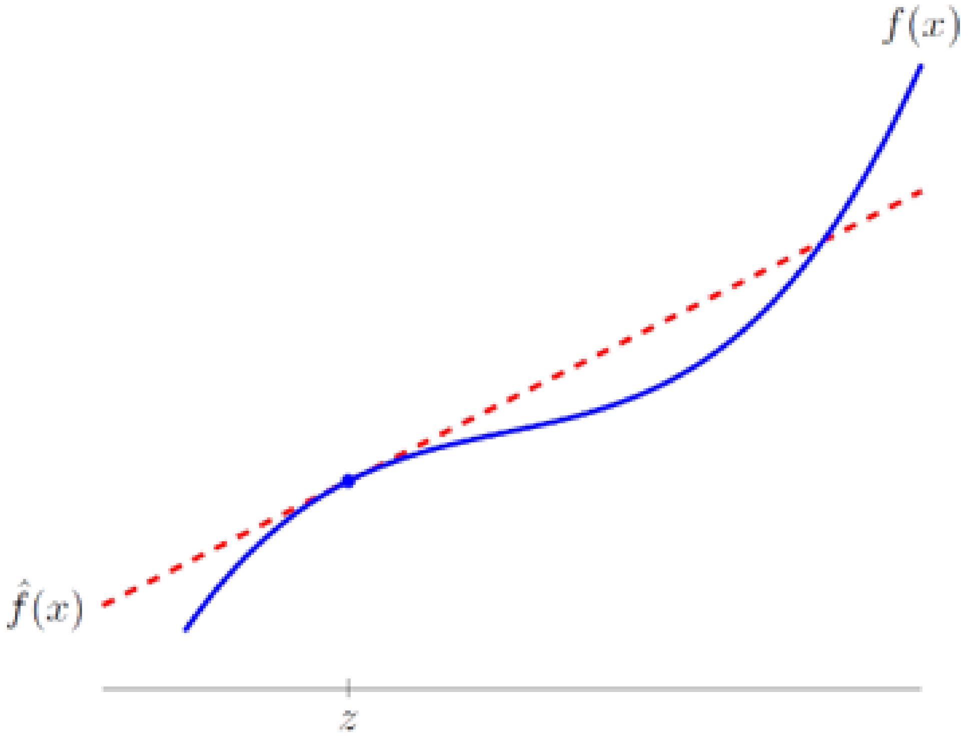
- $\hat{f}(x)$ is a linear function or a affine function?

$$\hat{f}(x) = \underbrace{f(z)}_{\text{Constant- value of function at z}} + \nabla f(z)^T \underbrace{(x - z)}_{\text{Deviation or Perturbation of x from z}}$$

$$\hat{f}(x) = \underbrace{\nabla f(z)^T x}_{\text{Linear function}} + \underbrace{(f(z) - \nabla f(z)^T z)}_{\text{Constant}}$$

Taylor approximation

- The Taylor approximation is sometimes called the linear approximation or linearized approximation of f (at z)



A function f of one variable, and the first order Taylor approximation $\hat{f}(x) = f(z) + f'(z)(x - z)$ at z

Example

Consider the function $f\colon \mathcal{R}^2 \to \mathcal{R}$ give by $f(x) = x_1 + \exp(x_2 - x_1)$, the Taylor approximation $\hat{f}$ near the point $z = (1,2)$

$e \sim 2.71$

$e^{-1} \sim 0.367$

$\boldsymbol{x}$	$\boldsymbol{f(x)}$	$\boldsymbol{\hat{f}(x)}$	$\boldsymbol{\lvert\hat{f}(x) - f(x)\rvert}$
(1.00,2.00)	3.7183	3.7183	0.0000
(0.96,1.98)	3.7332	3.7326	0.0005
(1.10,2.11)	3.8456	3.8455	0.0001
(0.85,2.05)	4.1701	4.1119	0.0582
(1.25,2.41)	4.4399	4.4032	0.0367

Definition

- Regression model is (the affine function of x):

$$\hat{y} = x^T w + w_0$$

$$\hat{y} = x^T w$$

Regression model

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 :

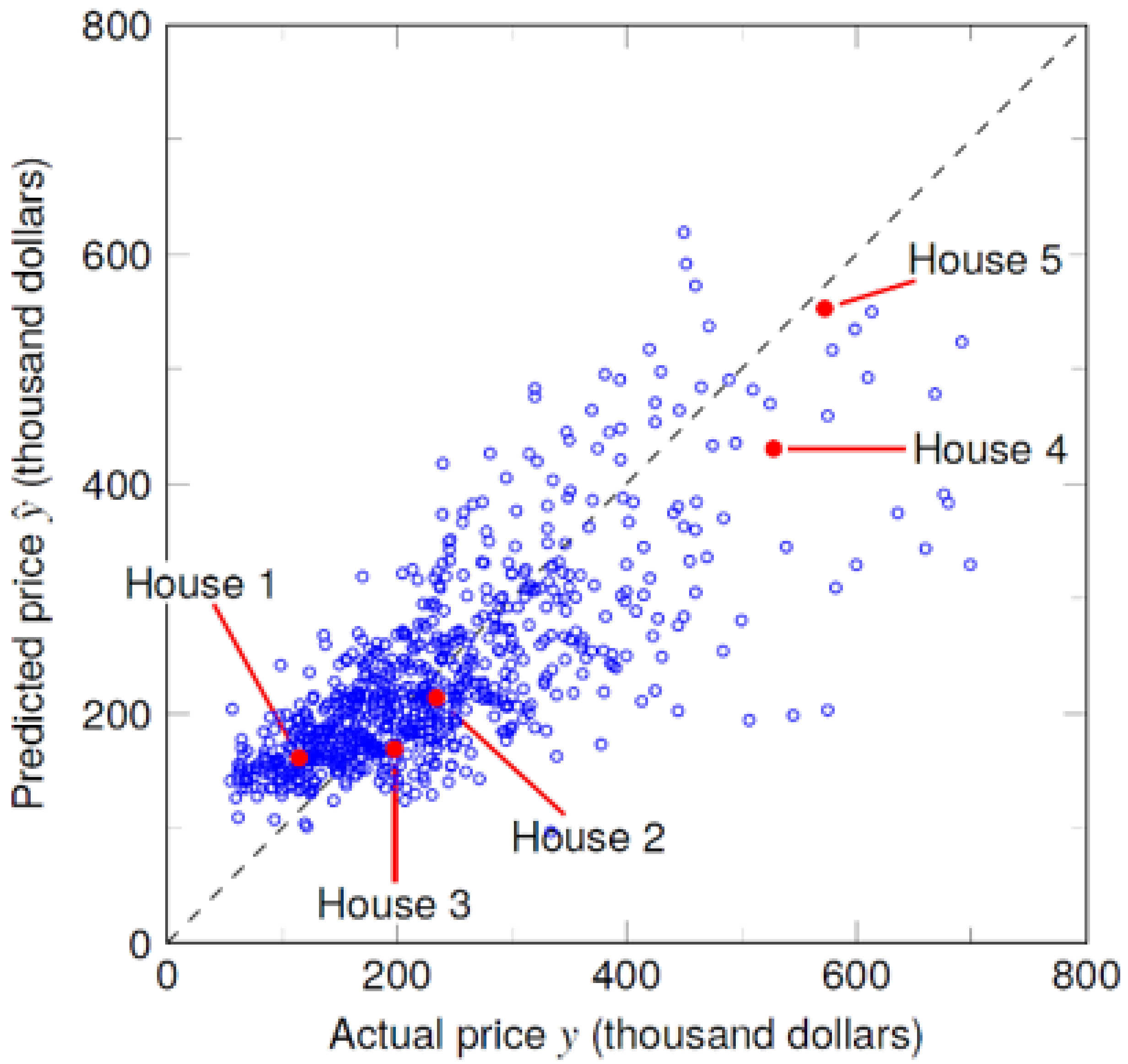
$$\beta = (148.73, -18.85), \qquad v = 54.40$$

- We'll see later how to guess β and v from sales data

Regression model

$House$	x_1 (area)	x_2 (beds)	y (prince)	$\hat{y}$(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

Regression model



Example

- What happened when feature is zero vector?
- Find the age based on following features:
  - What are the constraints?

Gender		Diabetes		Smoking		Age
Female	Male	Yes	No	Yes	No	

Reference

- Chapter 2: Introduction to Applied Linear Algebra Vectors, Matrices, and Least Squares
- Part of chapter 1 and chapter 6: Linear Algebra by David Cherney, etc.
- http://vmls-book.Stanford.edu/vmls-slides.pdf