Optimization of Communication Systems
EE 669

Chathuranga Weeraddana

March 2022

Lecture 4: Functions

Functions

Definition

- notations:
$$f : \mathcal{X} \to \mathcal{Y}$$
- f is a rule which assigns to x in the set $\mathcal{X}$ a unique y in the set $\mathcal{Y}$
- $\mathcal{X}$, the domain of f: the set of possible input values for f
- $\{f(x) \mid x \in \mathcal{X}\}$, the range of f: the set of all possible output values for f
- $\mathcal{Y}$, the codomain of f: a sort of target for the outputs[1]
- e.g., $f(x) = 1/x$

[1] see Defintion 2.2 Hammack-2009

Graph of f

- graph of $f : \mathcal{X} \to \mathcal{Y}$ is

$$\mathcal{G}(f) = \{(x, f(x) \mid x \in \mathcal{X}\} = \{(x, f(x) \mid x \in \mathtt{dom}(f)\}$$

- e.g., sketch the graphs of
  - $\sin x$
  - $1/x$
  - $\log x$
  - e^x

Elementary Functions

- power function
- exponential functions
- trigonometric functions [2]

[2] See § 2.3 of Binmore-2012

Functions of Several Variables

Real valued functions of two variables

- $f : \mathbb{R}^2 \to \mathbb{R}$
- f is a rule which assigns to $\boldsymbol{x} = (x_1, x_2)$ in the set $\mathbb{R}^2$ a unique x_3 in the set $\mathbb{R}$
- choose to think of it as a landscape
- how?
  - drawing a map of horizontal curves at a fixed level
  - called contours or level sets

Real valued functions of two variables

- e.g., graphs of (illustrate with Geogebra)
  - $x_1^2 + x_2^2$
  - $1/x_1$
  - $x_1^2 - 4x_2$
  - $x_1 + 2x_2$

Linear and Affine Functions On $\mathbb{R}^2$

- $f(\boldsymbol{x}) = a_1 x_1 + a_2 x_2 + c_2$
- identify the $\mathcal{G}(f)$ and level sets
- see § 3.1.1 of of Binmore-2012

Quadratic Functions On $\mathbb{R}^2$

- $f(\boldsymbol{x}) = \boldsymbol{x}^\mathsf{T} P \boldsymbol{x} + \boldsymbol{q}^\mathsf{T} \boldsymbol{x} + r$
- identify the $\mathcal{G}(f)$ and level sets
- see § 3.1.2 of of Binmore-2012

Derivative

- derivative of scalar valued functions defined on $\mathbb{R}$
- definition
- differentiability at a point $x = x_0$
- differentiability in an interval

Derivative

- derivative of scalar valued functions defined on $\mathbb{R}^n$
- definition
- differentiability at a point $\mathbf{x} = \mathbf{x}_0$
- differentiability in a set $\mathcal{X} \in \mathbb{R}^n$

Derivative

- suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ and x is an element in $\mathbb{R}^n$ and strictly inside the domain f
- f is differentiable at $\mathbf{c}$ if there exist a $1 \times n$ matrix $Df(\mathbf{c})$ that satisfy

$$\lim_{||\mathbf{x}-\mathbf{c}||_2 \to 0} \frac{||f(\mathbf{x}) - f(\mathbf{c}) - Df(\mathbf{c})(\mathbf{x}-\mathbf{c})||_2}{||\mathbf{x}-\mathbf{c}||_2} = 0 \qquad (1)$$

- the matrix $Df(\mathbf{c})$ is called the derivative of f at $\mathbf{c}$

Derivative

- the derivative can be found by deriving the matrix $Df(\mathbf{c})$ that satisfies (1)
- the derivative can also be found from partial derivatives

$$Df(\mathbf{x})_{1,j} = \frac{\partial f(\mathbf{x})}{\partial x_j},$$

  provided that the partial derivative exists
- note that $j \in \{1, 2, \ldots, n\}$

Derivative

- from a geometric point of view, the existence of the derivative of f at $\mathbf{c}$ reflects the existence of a tangent plane to the graph of f at the point $(\mathbf{c}, f(\mathbf{c}))$

Derivative

- some notations !
- recall that the derivative of f at $\mathbf{x}$ is $Df(\mathbf{x})$
- similarly, the derivative of f at $\mathbf{c}$ is $Df(\mathbf{c})$
- when we are given the derivative at a point $\mathbf{x}$, [i.e., $Df(\mathbf{x})$], we can easily compute the derivative at another point $\mathbf{c}$ as follows:

$$Df(\mathbf{c}) = Df(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{c}}$$

First-Order Approximation of f at x

- consider the function f_1 of $\mathbf{x}$ given by

$$f_1(\mathbf{x}) = f(\mathbf{c}) + Df(\mathbf{c})(\mathbf{x} - \mathbf{c})$$

- f_1 is called the first-order approximation of f at (or near) $\mathbf{c}$
- when $\mathbf{x}$ is close to $\mathbf{c}$, f_1 is very close to f
- graphically interpret this phenomenon

Gradient

- gradient of a real-valued function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ at $\mathbf{x}$ is

$$\nabla f(\mathbf{x}) = Df(\mathbf{x})^T$$

- i.e., the transpose of the derivative

- note that $\nabla f(\mathbf{x}) \in \mathbb{R}^n$, a n-vector

- moreover, it is easy to see that

$$\nabla f(\mathbf{x})_j = \frac{\partial f(\mathbf{x})}{\partial x_j}, \quad j = 1, \ldots, n$$

- the first-order approximation f_1 of f at $\mathbf{c}$ is simply

$$f_1(\mathbf{x}) = f(\mathbf{c}) + \nabla f(\mathbf{c})^T(\mathbf{x} - \mathbf{c})$$

Gradient

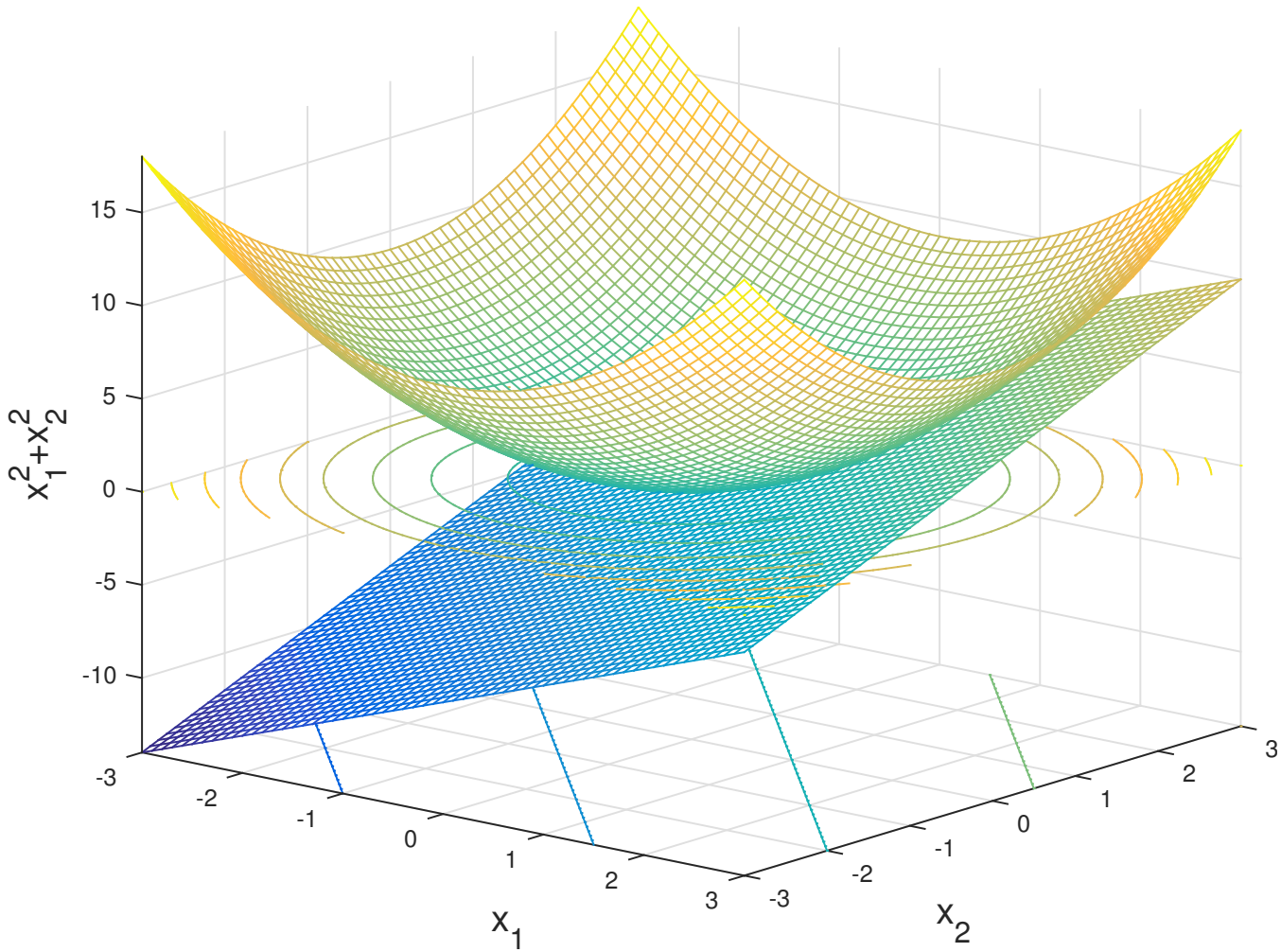
- determine the derivatives, gradients, and first-order approximations of
  - $f(\mathbf{x}) = ||\mathbf{x}||^2$
  - $f(x_1, x_2) = x_1^2 + 2x_2$
  - $f(x_1, x_2) = \sin(x_1) + x_2$
  - $f(x_1, x_2) = 2x_1 + 3x_2$

  at $\mathbf{c}$

- note that

$$\nabla f(\mathbf{c}) = \nabla f(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{c}}$$

First-Order Approximation of $f(\mathbf{x}) = ||\mathbf{x}||^2$ at $\mathbf{x} = (1, 1)$



Derivative, Gradient

- some useful results
  - $D(\mathbf{q}^T\mathbf{x}) = \mathbf{q}^T, \quad \mathbf{q} \in \mathbb{R}^m$
  - $D(\mathbf{q}^T A\mathbf{x}) = \mathbf{q}^T A, \quad \mathbf{q} \in \mathbb{R}^m, A \in \mathbb{R}^{m\times n}$
  - $D(\mathbf{x}^T A\mathbf{x}) = \mathbf{x}^T(A + A^T), \quad A \in \mathbb{R}^{n\times n}$
  - $D(\mathbf{x}^T A\mathbf{x}) = \mathbf{x}^T(A + A^T) = 2\mathbf{x}^T A, \quad A \in \mathbb{S}^{n\times n}$
  - $D(\mathbf{x}^T\mathbf{x}) = \mathbf{x}^T(I + I^T) = 2\mathbf{x}^T I = 2\mathbf{x}^T$
- recall: the gradient is simply the transpose of the derivative

Second Derivative

- suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $\mathbf{x}$ is an element in $\mathbb{R}^n$ and in the interior of domain f
- the second derivative of Hessian matrix of f at $\mathbf{x}$, denoted $\nabla^2 f(\mathbf{x})$, is given by

$$\nabla^2 f(\mathbf{x})_{i,j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, n$$

provided f is twice differentiable at $\mathbf{x}$

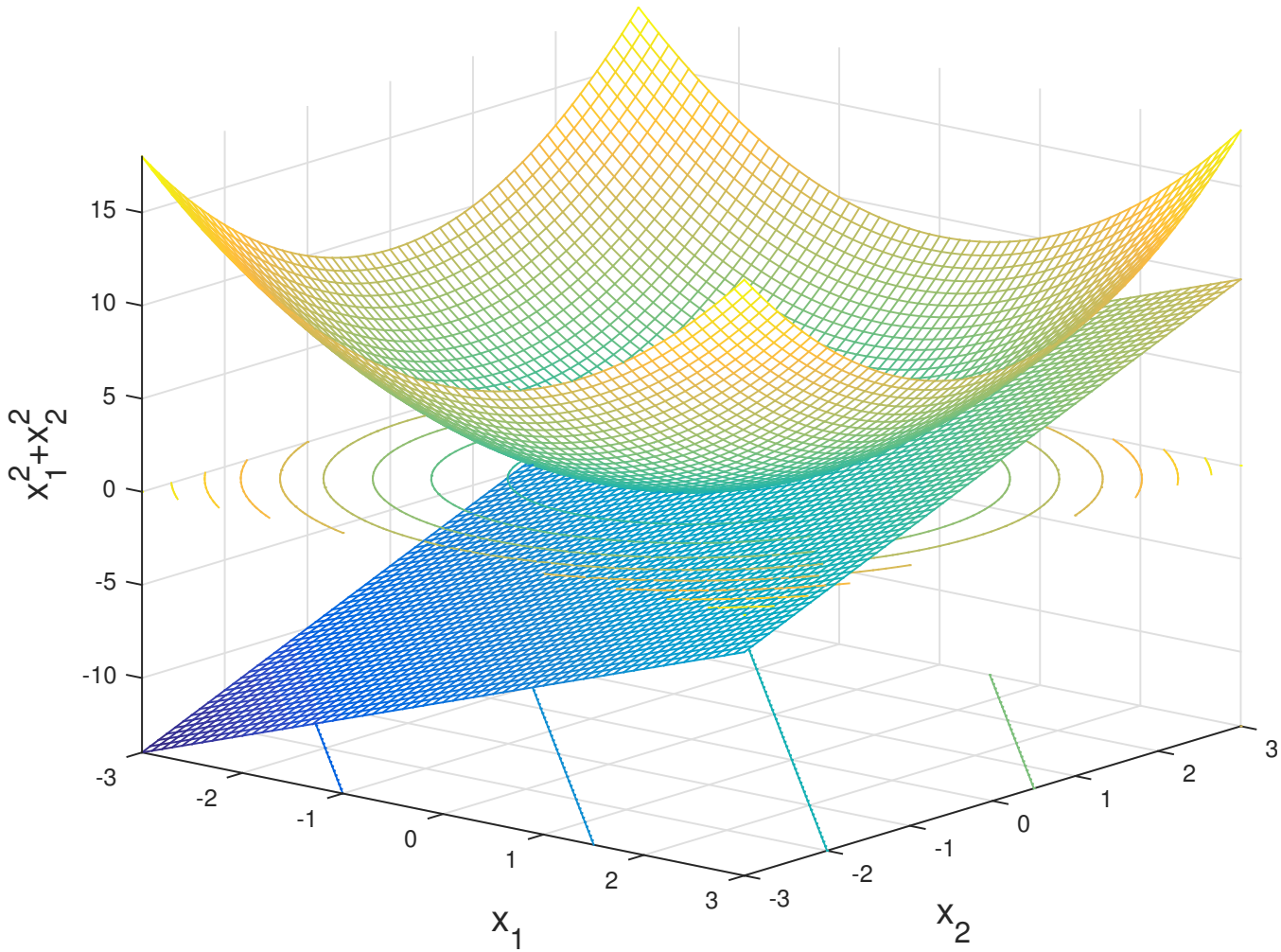
Second-Order Approximation of f at $\mathbf{x}$

- consider the function f_2 of $\mathbf{x}$ given by

$$f_1(\mathbf{x}) = f(\mathbf{c}) + \nabla f(\mathbf{c})^T(\mathbf{x}-\mathbf{c}) + (1/2)(\mathbf{x}-\mathbf{c})^T\nabla^2 f(\mathbf{c})(\mathbf{x}-\mathbf{c})$$

- f_2 is called the second-order approximation of f at (or near) $\mathbf{c}$
- when $\mathbf{x}$ is close to $\mathbf{c}$, f_2 is very very close to f
- graphically interpret this phenomenon

Second-Order Approximation of $f(x) = ||x||^2$ at $x = (1, 1)$



Second Derivative

- determine the Hessian and second-order approximations of
  - $f(\mathbf{x}) = ||\mathbf{x}||^2$
  - $f(x_1, x_2) = x_1^2 + 2x_2$
  - $f(x_1, x_2) = \sin(x_1) + x_2$
  - $f(x_1, x_2) = 2x_1 + 3x_2$

  at $\mathbf{c}$
- note that

$$\nabla^2 f(\mathbf{c}) = \nabla^2 f(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{c}}$$

Second Derivative

- some useful results
  - $\nabla^2(\mathbf{q}^T\mathbf{x}) = 0 \in \mathbb{R}^{m\times m}, \quad \mathbf{q} \in \mathbb{R}^m$
  - $\nabla^2(\mathbf{q}^T A\mathbf{x}) = 0 \in \mathbb{R}^{n\times n}, \quad \mathbf{q} \in \mathbb{R}^m, A \in \mathbb{R}^{m\times n}$
  - $\nabla^2(\mathbf{x}^T A\mathbf{x}) = (A + A^T), \quad A \in \mathbb{R}^{n\times n}$
  - $\nabla^2(\mathbf{x}^T A\mathbf{x}) = (A + A^T) = 2A, \quad A \in \mathbb{S}^{n\times n}$
  - $\nabla^2(\mathbf{x}^T\mathbf{x}) = (I + I^T) = 2I$

Conics

Conics

- we have defined functions expressed explicitly in terms of an independent variable
- functions may also be defined implicitly by a relation between x and y
- a simple one is a second degree equation:

$$ax^2 + by^2 + cxy + dx + ey + f = 0. \tag{2}$$

- the above always represents a conic section, the intersection of a doublenapped cone and a plane
- discuss: [see § 2.14 of Binmore-2012]